



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 04:06 PM UTC

PDB ID : 3FOA / pdb\_00003foa  
Title : Crystal structure of the bacteriophage T4 tail sheath protein, deletion mutant gp18M  
Authors : Aksyuk, A.A.; Leiman, P.G.; Kurochkina, L.P.; Shneider, M.M.; Kostyuchenko, V.A.; Mesyanzhinov, V.V.; Rossmann, M.G.  
Deposited on : 2008-12-29  
Resolution : 3.50 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  
Xtriage (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

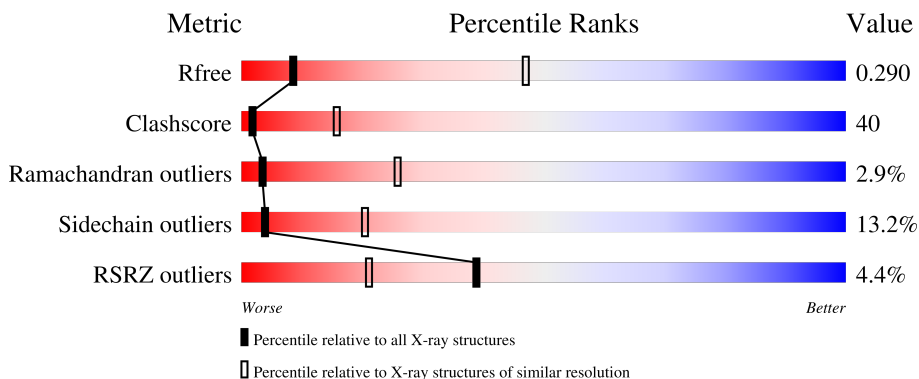
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	180053	1085 (3.54-3.46)
Clashscore	190562	1140 (3.54-3.46)
Ramachandran outliers	187476	1113 (3.54-3.46)
Sidechain outliers	187428	1114 (3.54-3.46)
RSRZ outliers	180081	1084 (3.54-3.46)

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	510	
1	B	510	
1	C	510	
1	D	510	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 14406 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 sheath protein Gp18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	479	Total 3613	C 2279	N 600	O 726	S 8	0	0	0
1	B	478	Total 3605	C 2275	N 598	O 724	S 8	0	0	0
1	C	477	Total 3597	C 2271	N 596	O 722	S 8	0	0	0
1	D	476	Total 3591	C 2268	N 595	O 720	S 8	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	GLU	ASP	variant	UNP P13332
A	148	ALA	GLY	variant	UNP P13332
A	150	ILE	ASN	variant	UNP P13332
A	151	ILE	TYR	variant	UNP P13332
A	301	GLY	GLU	variant	UNP P13332
A	399	VAL	ALA	variant	UNP P13332
A	454	TYR	HIS	variant	UNP P13332
A	510	PRO	ARG	engineered mutation	UNP P13332
B	100	GLU	ASP	variant	UNP P13332
B	148	ALA	GLY	variant	UNP P13332
B	150	ILE	ASN	variant	UNP P13332
B	151	ILE	TYR	variant	UNP P13332
B	301	GLY	GLU	variant	UNP P13332
B	399	VAL	ALA	variant	UNP P13332
B	454	TYR	HIS	variant	UNP P13332
B	510	PRO	ARG	engineered mutation	UNP P13332
C	100	GLU	ASP	variant	UNP P13332
C	148	ALA	GLY	variant	UNP P13332
C	150	ILE	ASN	variant	UNP P13332
C	151	ILE	TYR	variant	UNP P13332
C	301	GLY	GLU	variant	UNP P13332

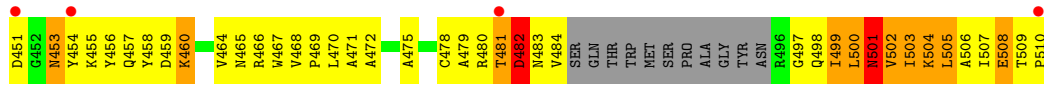
*Continued on next page...*

*Continued from previous page...*

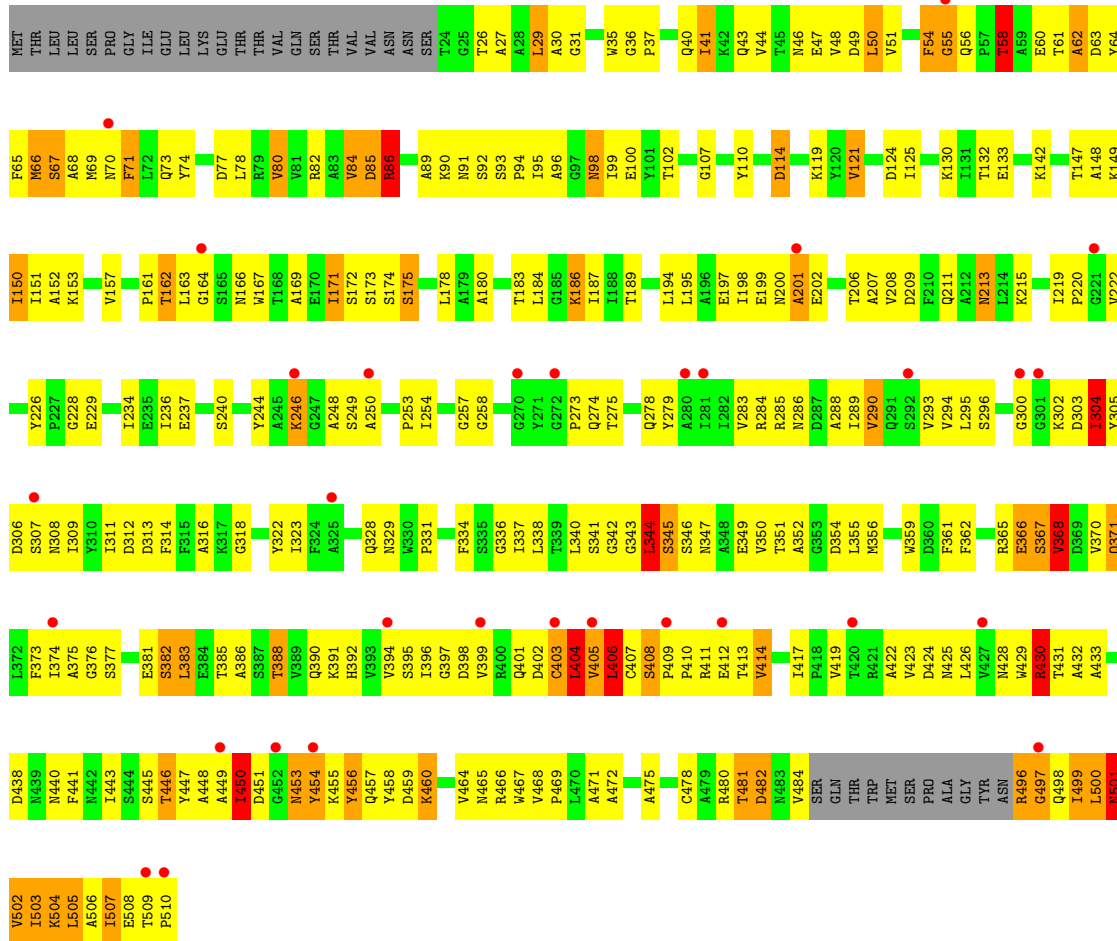
<b>Chain</b>	<b>Residue</b>	<b>Modelled</b>	<b>Actual</b>	<b>Comment</b>	<b>Reference</b>
C	399	VAL	ALA	variant	UNP P13332
C	454	TYR	HIS	variant	UNP P13332
C	510	PRO	ARG	engineered mutation	UNP P13332
D	100	GLU	ASP	variant	UNP P13332
D	148	ALA	GLY	variant	UNP P13332
D	150	ILE	ASN	variant	UNP P13332
D	151	ILE	TYR	variant	UNP P13332
D	301	GLY	GLU	variant	UNP P13332
D	399	VAL	ALA	variant	UNP P13332
D	454	TYR	HIS	variant	UNP P13332
D	510	PRO	ARG	engineered mutation	UNP P13332







• Molecule 1: Tail sheath protein Gp18



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.59Å 116.29Å 433.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.80 – 3.50 49.80 – 3.50	Depositor EDS
% Data completeness (in resolution range)	87.8 (49.80-3.50) 87.8 (49.80-3.50)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.68 (at 3.48Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.267 , 0.299 0.281 , 0.290	Depositor DCC
$R_{free}$ test set	5405 reflections (16.72%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.5	Xtrriage
Anisotropy	0.673	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 71.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	14406	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% 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

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	0.47	4/3677 (0.1%)	1.09	30/5001 (0.6%)
1	B	0.38	1/3669 (0.0%)	1.20	45/4990 (0.9%)
1	C	0.40	0/3661	1.08	35/4979 (0.7%)
1	D	0.38	0/3655	1.29	48/4971 (1.0%)
All	All	0.41	5/14662 (0.0%)	1.17	158/19941 (0.8%)

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	13
1	C	0	5
1	D	0	2
All	All	0	24

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	22	ASN	C-N	-7.18	1.23	1.33
1	A	23	SER	C-N	-6.90	1.24	1.33
1	A	23	SER	C-O	6.82	1.32	1.24
1	B	432	ALA	C-N	6.65	1.42	1.33
1	A	22	ASN	C-O	5.52	1.30	1.24

The worst 5 of 158 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	55	GLY	N-CA-C	33.86	153.61	112.64
1	C	195	LEU	N-CA-C	21.34	134.23	110.97
1	B	96	ALA	N-CA-CB	-21.11	78.81	110.26

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	503	ILE	CB-CA-C	20.59	129.38	111.05
1	A	55	GLY	N-CA-C	18.96	158.11	113.18

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	21	ASN	Peptide
1	A	450	ILE	Peptide
1	A	451	ASP	Peptide
1	A	452	GLY	Peptide
1	B	22	ASN	Peptide

## 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	3613	0	3554	323	2
1	B	3605	0	3548	304	1
1	C	3597	0	3541	272	1
1	D	3591	0	3536	249	0
All	All	14406	0	14179	1135	2

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 40.

The worst 5 of 1135 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:A:409:PRO:O	1:A:454:TYR:CE1	1.65	1.48
1:A:409:PRO:CD	1:A:451:ASP:O	1.75	1.33
1:B:496:ARG:HG3	1:B:497:GLY:O	1.16	1.29
1:A:379:ALA:CB	1:A:454:TYR:OH	1.84	1.25
1:B:496:ARG:CG	1:B:497:GLY:O	1.89	1.21

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:509:THR:O	1:B:419:VAL:N[4_655]	1.92	0.28
1:A:22:ASN:ND2	1:C:260:THR:OG1[4_555]	2.00	0.20

## 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	475/510 (93%)	371 (78%)	87 (18%)	17 (4%)	2	22
1	B	474/510 (93%)	365 (77%)	94 (20%)	15 (3%)	3	24
1	C	473/510 (93%)	374 (79%)	84 (18%)	15 (3%)	3	24
1	D	472/510 (92%)	377 (80%)	87 (18%)	8 (2%)	7	35
All	All	1894/2040 (93%)	1487 (78%)	352 (19%)	55 (3%)	3	26

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	98	ASN
1	A	201	ALA
1	A	432	ALA
1	B	403	CYS
1	B	502	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 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	383/411 (93%)	335 (88%)	48 (12%)	4	22
1	B	382/411 (93%)	336 (88%)	46 (12%)	5	23
1	C	381/411 (93%)	329 (86%)	52 (14%)	3	19
1	D	380/411 (92%)	324 (85%)	56 (15%)	3	17
All	All	1526/1644 (93%)	1324 (87%)	202 (13%)	4	20

5 of 202 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	213	ASN
1	C	499	ILE
1	D	502	VAL
1	C	341	SER
1	C	406	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	328	GLN
1	C	483	ASN
1	C	329	ASN
1	C	439	ASN
1	D	43	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](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 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	479/510 (93%)	0.35	12 (2%) 58 35	24, 65, 121, 227	0
1	B	478/510 (93%)	0.66	28 (5%) 28 16	20, 82, 162, 247	0
1	C	477/510 (93%)	0.42	13 (2%) 56 33	25, 80, 148, 263	0
1	D	476/510 (93%)	0.76	31 (6%) 25 15	26, 116, 254, 407	0
All	All	1910/2040 (93%)	0.55	84 (4%) 39 21	20, 84, 186, 407	0

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	454	TYR	5.9
1	B	454	TYR	5.6
1	B	48	VAL	5.5
1	A	510	PRO	5.4
1	A	22	ASN	4.9

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.