



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 03:05 PM UTC

PDB ID : 9T4P / pdb\_00009t4p  
Title : Dengue 2 Virus NS2B-NS3 protease fusion protein with crystal epitope mutation K174Q  
Authors : Fairhead, M.; Ni, X.; Koekemoer, L.; MacLean, E.M.; Wright, N.D.; Bowsman-Jones, H.; Von Delft, F.  
Deposited on : 2025-10-30  
Resolution : 1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

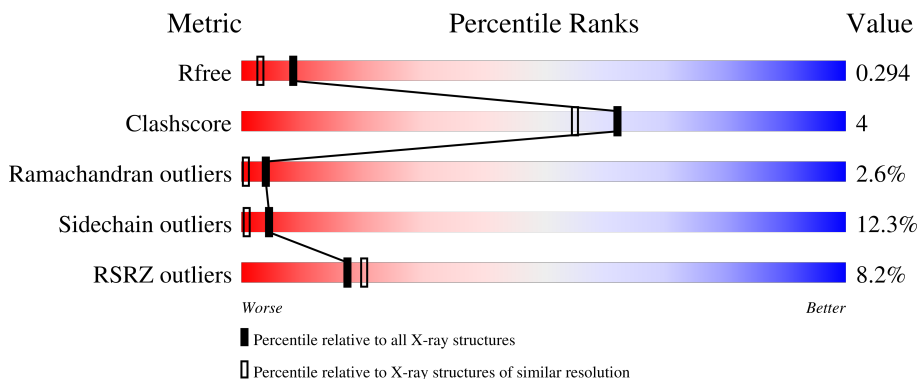
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	3494 (1.96-1.96)
Clashscore	190562	3612 (1.96-1.96)
Ramachandran outliers	187476	3587 (1.96-1.96)
Sidechain outliers	187428	3587 (1.96-1.96)
RSRZ outliers	180081	3495 (1.96-1.96)

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	244	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2966 atoms, of which 1435 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 Serine protease subunit NS2B,Serine protease NS3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	196	2918	933	1435	253	294	3	0	0	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ALA	-	expression tag	UNP P14337
A	1	SER	-	expression tag	UNP P14337
A	2	MET	-	expression tag	UNP P14337
A	50	GLY	-	linker	UNP P14337
A	51	GLY	-	linker	UNP P14337
A	52	GLY	-	linker	UNP P14337
A	53	GLY	-	linker	UNP P14337
A	54	SER	-	linker	UNP P14337
A	55	GLY	-	linker	UNP P14337
A	56	GLY	-	linker	UNP P14337
A	57	GLY	-	linker	UNP P14337
A	58	GLY	-	linker	UNP P14337
A	175	GLN	LYS	engineered mutation	UNP P14337

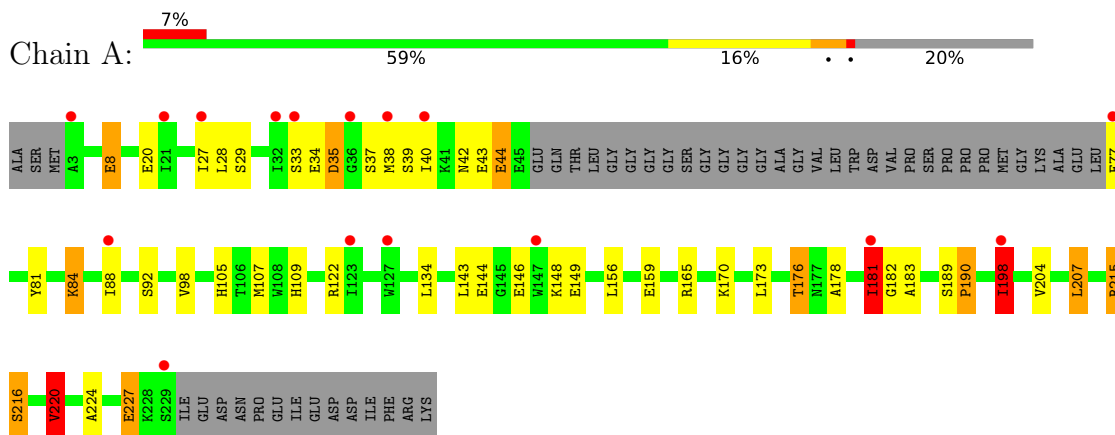
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	48	Total	O	0	0
			48	48		

### 3 Residue-property plots [i](#)

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: Serine protease subunit NS2B,Serine protease NS3



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	123.49Å 45.20Å 40.46Å 90.00° 109.11° 90.00°	Depositor
Resolution (Å)	42.15 – 1.95 42.15 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.15-1.95) 99.9 (42.15-1.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.78 (at 1.95Å)	Xtrriage
Refinement program	REFMAC 5.8.0431 (refmacat 0.4.105)	Depositor
R, $R_{free}$	0.226 , 0.294 0.226 , 0.294	Depositor DCC
$R_{free}$ test set	800 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.0	Xtrriage
Anisotropy	0.318	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 48.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.042 for -h-2*1,-k,l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2966	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.58% 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	0.98	2/1510 (0.1%)	1.68	27/2035 (1.3%)

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	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	109	HIS	CE1-NE2	5.50	1.38	1.32
1	A	105	HIS	CE1-NE2	-5.11	1.27	1.32

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	8	GLU	CG-CD-OE2	9.80	140.94	118.40
1	A	20	GLU	CB-CA-C	-8.05	97.92	110.81
1	A	220	VAL	N-CA-CB	-7.85	97.84	112.36
1	A	20	GLU	N-CA-CB	7.44	121.02	109.94
1	A	34	GLU	CB-CG-CD	7.39	125.17	112.60
1	A	8	GLU	CG-CD-OE1	-7.09	102.09	118.40
1	A	216	SER	N-CA-CB	7.04	122.39	110.49
1	A	37	SER	CA-CB-OG	6.86	124.81	111.10
1	A	227	GLU	N-CA-CB	6.38	119.99	110.29
1	A	149	GLU	CB-CA-C	6.24	119.88	109.89
1	A	170	LYS	CB-CA-C	6.18	117.03	110.17
1	A	159	GLU	N-CA-CB	5.92	119.36	110.02
1	A	44	GLU	N-CA-CB	5.89	120.44	110.49
1	A	198	ILE	CB-CA-C	5.75	119.14	110.63

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	28	LEU	CB-CA-C	5.72	119.18	109.75
1	A	84	LYS	CD-CE-NZ	5.48	129.43	111.90
1	A	181	ILE	CA-C-N	5.39	125.46	120.34
1	A	181	ILE	C-N-CA	5.39	125.46	120.34
1	A	35	ASP	CA-C-N	5.38	125.45	120.34
1	A	35	ASP	C-N-CA	5.38	125.45	120.34
1	A	84	LYS	CB-CA-C	5.33	118.98	109.38
1	A	198	ILE	CA-CB-CG2	5.29	119.49	110.50
1	A	144	GLU	CB-CA-C	5.28	118.82	109.75
1	A	148	LYS	CB-CA-C	5.19	118.73	109.38
1	A	144	GLU	N-CA-CB	-5.08	103.64	110.95
1	A	189	SER	CA-C-N	5.04	126.14	119.84
1	A	189	SER	C-N-CA	5.04	126.14	119.84

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	122	ARG	Sidechain
1	A	165	ARG	Sidechain
1	A	215	ARG	Sidechain

## 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	1483	1435	1469	12	0
2	A	48	0	0	1	0
All	All	1531	1435	1469	12	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 4.

All (12) 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:8:GLU:OE2	1:A:84:LYS:NZ	2.11	0.81
1:A:215:ARG:O	2:A:301:HOH:O	2.09	0.68
1:A:107:MET:HE2	1:A:107:MET:HA	1.82	0.62
1:A:107:MET:HE3	1:A:207:LEU:HD23	1.84	0.59
1:A:183:ALA:HB1	1:A:220:VAL:HG13	1.93	0.49
1:A:134:LEU:HD21	1:A:224:ALA:HB1	1.96	0.48
1:A:176:THR:HG23	1:A:178:ALA:H	1.78	0.47
1:A:29:SER:HB2	1:A:40:ILE:HG12	1.96	0.47
1:A:77:GLU:O	1:A:81:TYR:OH	2.32	0.46
1:A:107:MET:HE2	1:A:107:MET:CA	2.47	0.43
1:A:198:ILE:HD11	1:A:204:VAL:HG22	2.00	0.43
1:A:181:ILE:HG13	1:A:182:GLY:O	2.19	0.43

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 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	192/244 (79%)	181 (94%)	6 (3%)	5 (3%)	<b>4</b> <b>1</b>

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	ASP
1	A	44	GLU
1	A	216	SER
1	A	43	GLU
1	A	190	PRO

### 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	155/190 (82%)	136 (88%)	19 (12%)	<b>4</b> <b>1</b>

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	ILE
1	A	33	SER
1	A	38	MET
1	A	39	SER
1	A	42	ASN
1	A	88	ILE
1	A	92	SER
1	A	98	VAL
1	A	143	LEU
1	A	146	GLU
1	A	156	LEU
1	A	173	LEU
1	A	176	THR
1	A	181	ILE
1	A	190	PRO
1	A	198	ILE
1	A	207	LEU
1	A	220	VAL
1	A	227	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	118	HIS

### 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	196/244 (80%)	0.46	16 (8%) <b>17</b> <b>20</b>	26, 49, 83, 106	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	40	ILE	5.6
1	A	32	ILE	4.9
1	A	3	ALA	3.4
1	A	38	MET	3.2
1	A	27	ILE	3.1
1	A	181	ILE	3.0
1	A	88	ILE	3.0
1	A	229	SER	2.7
1	A	123	ILE	2.6
1	A	198	ILE	2.4
1	A	33	SER	2.4
1	A	21	ILE	2.2
1	A	36	GLY	2.2
1	A	127	TRP	2.1
1	A	77	GLU	2.1
1	A	147	TRP	2.0

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.