



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 22, 2025 – 01:21 PM JST

PDB ID : 9IW2 / pdb\_00009iw2  
Title : Chikungunya virus E protein complexed with C37 Fab  
Authors : Qi, J.; Han, X.; Wang, F.; Tian, S.; Gao, F.G.; Yan, J.  
Deposited on : 2024-07-25  
Resolution : 3.20 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 Phenix2.0rc1  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.44

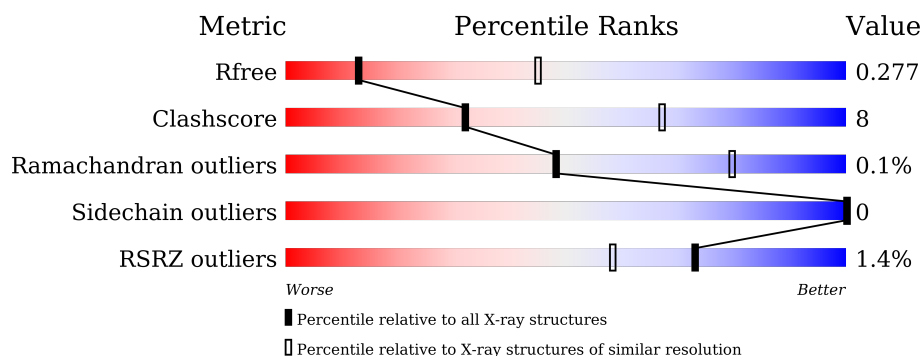
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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))
$R_{free}$	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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	F	432	<div> <div>3%</div> <div> <div></div> <div>67%</div> <div>24%</div> <div>9%</div> </div> </div>
2	H	265	<div> <div>%</div> <div> <div></div> <div>71%</div> <div>15%</div> <div>14%</div> </div> </div>
3	L	237	<div> <div></div> <div> <div></div> <div>76%</div> <div>14%</div> <div>11%</div> </div> </div>
4	P	406	<div> <div></div> <div> <div></div> <div>69%</div> <div>15%</div> <div>17%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9007 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 CHIKV E1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	393	Total	C	N	O	S	0	0	0
			2989	1890	501	574	24			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	GLY	-	expression tag	UNP A4L787
F	-18	GLY	-	expression tag	UNP A4L787
F	-17	GLY	-	expression tag	UNP A4L787
F	-16	GLY	-	expression tag	UNP A4L787
F	-15	SER	-	expression tag	UNP A4L787
F	-14	GLY	-	expression tag	UNP A4L787
F	-13	GLY	-	expression tag	UNP A4L787
F	-12	GLY	-	expression tag	UNP A4L787
F	-11	GLY	-	expression tag	UNP A4L787
F	-10	SER	-	expression tag	UNP A4L787
F	-9	GLY	-	expression tag	UNP A4L787
F	-8	GLY	-	expression tag	UNP A4L787
F	-7	GLY	-	expression tag	UNP A4L787
F	-6	GLY	-	expression tag	UNP A4L787
F	-5	SER	-	expression tag	UNP A4L787
F	-4	GLY	-	expression tag	UNP A4L787
F	-3	GLY	-	expression tag	UNP A4L787
F	-2	GLY	-	expression tag	UNP A4L787
F	-1	GLY	-	expression tag	UNP A4L787
F	0	SER	-	expression tag	UNP A4L787

- Molecule 2 is a protein called C37 Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	228	Total	C	N	O	S	0	0	0
			1715	1084	286	340	5			

- Molecule 3 is a protein called C37 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	212	Total	C	N	O	S	0	0	0
			1634	1020	281	328	5			

- Molecule 4 is a protein called Togavirin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	P	338	Total	C	N	O	S	0	0	0
			2669	1664	483	502	20			

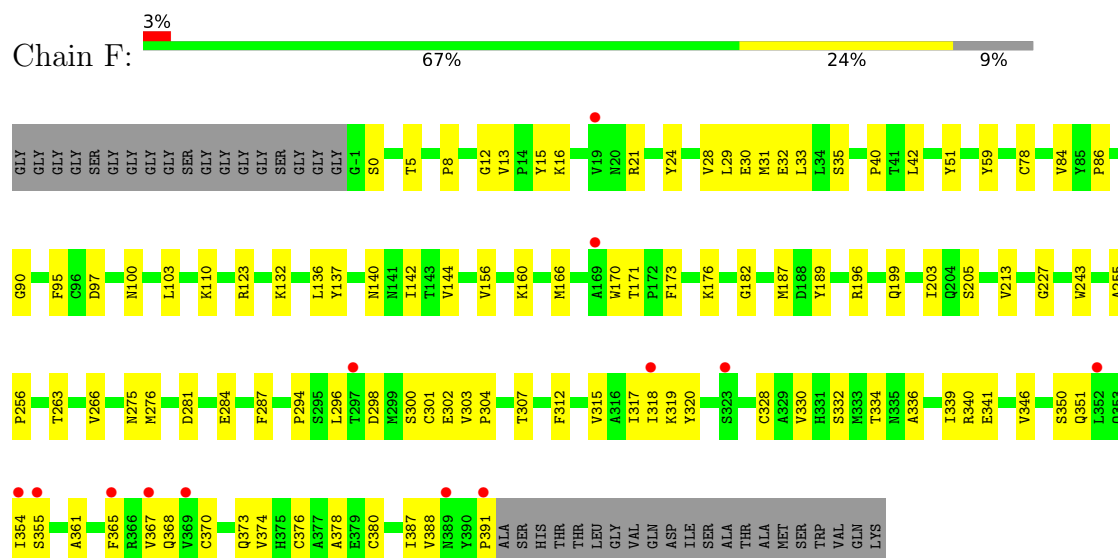
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	0	MET	-	expression tag	UNP C8YZ73

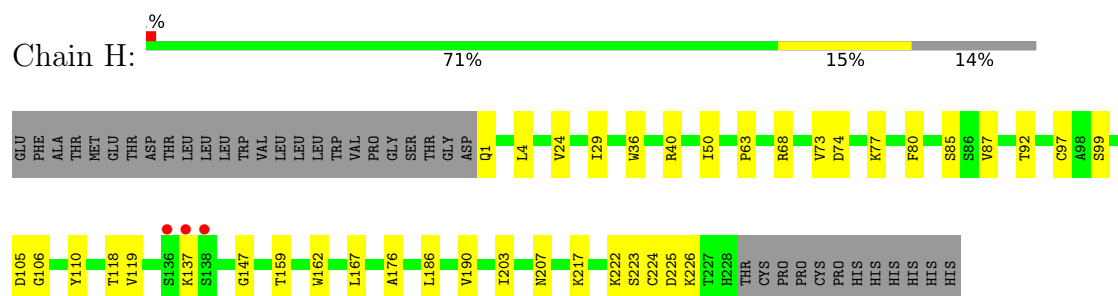
### 3 Residue-property plots [i](#)

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.

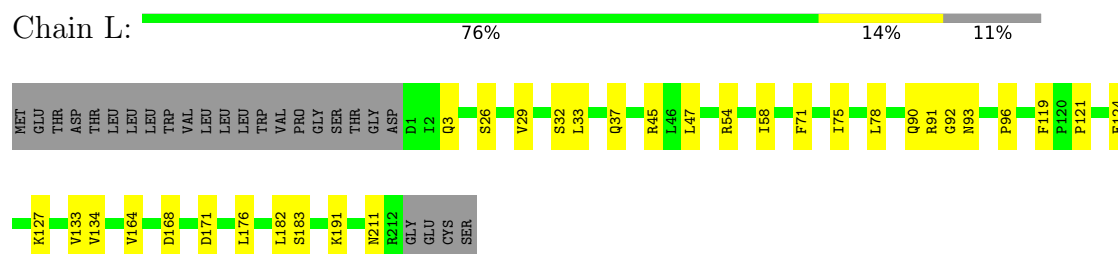
#### • Molecule 1: CHIKV E1



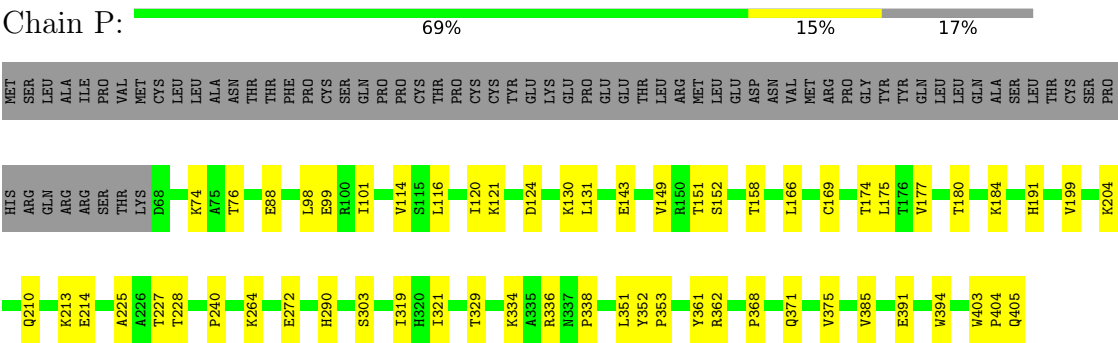
#### • Molecule 2: C37 Heavy Chain



#### • Molecule 3: C37 light chain



● Molecule 4: Togavirin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.09Å 119.00Å 228.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.85 – 3.20 46.85 – 3.20	Depositor EDS
% Data completeness (in resolution range)	91.2 (46.85-3.20) 91.2 (46.85-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487)	Depositor
R, $R_{free}$	0.222 , 0.277 0.222 , 0.277	Depositor DCC
$R_{free}$ test set	25278 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	73.6	Xtriage
Anisotropy	0.188	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 31.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	9007	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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	F	0.20	0/3066	0.51	0/4182
2	H	0.19	0/1763	0.52	0/2409
3	L	0.17	0/1671	0.45	0/2272
4	P	0.19	0/2741	0.53	0/3731
All	All	0.19	0/9241	0.51	0/12594

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
4	P	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	P	225	ALA	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	F	2989	0	2896	67	2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	1715	0	1664	23	0
3	L	1634	0	1584	21	0
4	P	2669	0	2582	36	2
All	All	9007	0	8726	140	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 8.

The worst 5 of 140 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:F:16:LYS:HZ1	1:F:340:ARG:HA	1.04	1.13
1:F:16:LYS:NZ	1:F:340:ARG:HA	1.87	0.89
1:F:33:LEU:HD21	1:F:276:MET:HE1	1.62	0.82
3:L:90:GLN:HE21	3:L:93:ASN:H	1.27	0.82
1:F:110:LYS:HG3	1:F:213:VAL:HG11	1.64	0.79

All (2) 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:F:123:ARG:NH2	4:P:88:GLU:OE2[1_455]	2.01	0.19
1:F:281:ASP:OD2	4:P:210:GLN:NE2[1_455]	2.08	0.12

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	F	391/432 (90%)	372 (95%)	18 (5%)	1 (0%)	37 69
2	H	226/265 (85%)	213 (94%)	13 (6%)	0	100 100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	L	210/237 (89%)	204 (97%)	6 (3%)	0	100	100
4	P	336/406 (83%)	316 (94%)	20 (6%)	0	100	100
All	All	1163/1340 (87%)	1105 (95%)	57 (5%)	1 (0%)	48	80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	0	SER

### 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	F	328/349 (94%)	328 (100%)	0	100	100
2	H	197/231 (85%)	197 (100%)	0	100	100
3	L	184/206 (89%)	184 (100%)	0	100	100
4	P	300/364 (82%)	300 (100%)	0	100	100
All	All	1009/1150 (88%)	1009 (100%)	0	100	100

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
4	P	82	HIS
4	P	211	HIS
4	P	345	ASN
4	P	296	HIS
3	L	138	ASN

### 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	F	393/432 (90%)	0.24	13 (3%) 49 34	46, 76, 229, 372	0
2	H	228/265 (86%)	-0.22	3 (1%) 74 60	44, 59, 118, 256	0
3	L	212/237 (89%)	-0.39	0 100 100	44, 60, 78, 91	0
4	P	338/406 (83%)	-0.32	0 100 100	44, 59, 91, 150	0
All	All	1171/1340 (87%)	-0.13	16 (1%) 73 58	44, 62, 201, 372	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	369	VAL	4.1
2	H	138	SER	3.7
1	F	352	LEU	2.6
1	F	391	PRO	2.6
1	F	389	ASN	2.5

### 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.