



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 3, 2024 – 08:39 pm BST

PDB ID : 6TVW  
Title : Structure of native gp41 derived peptide fusion inhibitor  
Authors : Huhmann, S.; Nyakatura, E.K.; Rohrhofer, A.; Schmidt, B.; Eichler, J.;  
Moschner, J.; Roth, C.; Kokschi, B.  
Deposited on : 2020-01-10  
Resolution : 1.45 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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.37.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.37.1

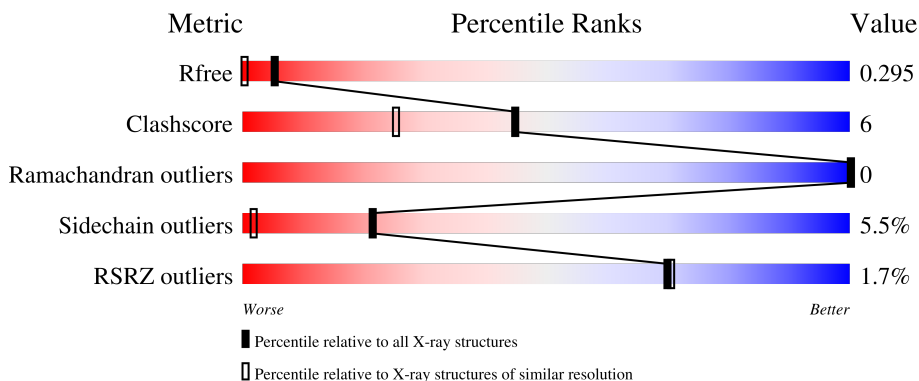
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.45 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))
$R_{free}$	130704	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

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	CCC	38	
2	DbD	31	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1171 atoms, of which 550 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 Envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
1	CCC	38	695	213	360	64	58	22	2	0

- Molecule 2 is a protein called Transmembrane protein gp41,Envelope glycoprotein gp160.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace	
			Total	C	F	H	N	O				S
2	DbD	23	402	132	3	190	33	42	2	12	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
DbD	701	O0B	-	expression tag	UNP P04582

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	CCC	58	Total	O	0	0
			58	58		
3	DbD	16	Total	O	0	0
			16	16		

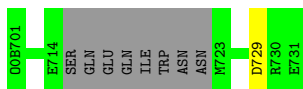
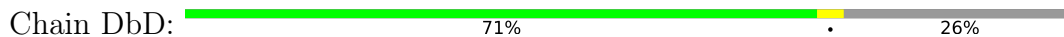
### 3 Residue-property plots [i](#)

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: Envelope glycoprotein



- Molecule 2: Transmembrane protein gp41,Envelope glycoprotein gp160



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	44.71Å 44.71Å 208.86Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	69.62 – 1.45 69.62 – 1.45	Depositor EDS
% Data completeness (in resolution range)	99.9 (69.62-1.45) 99.9 (69.62-1.45)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.68 (at 1.45Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.234 , 0.298 0.232 , 0.295	Depositor DCC
$R_{free}$ test set	690 reflections (4.66%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.9	Xtrriage
Anisotropy	0.630	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 68.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	1171	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 31.78 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.0461e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: O0B

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	CCC	0.66	0/338	0.76	0/453
2	DbD	0.73	0/205	0.77	0/276
All	All	0.69	0/543	0.76	0/729

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	CCC	335	360	354	4	0
2	DbD	212	190	178	0	0
3	CCC	58	0	0	4	0
3	DbD	16	0	0	0	0
All	All	621	550	532	4	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 6.

All (4) 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:CCC:580:ILE:HB	3:CCC:635:HOH:O	1.63	0.97
1:CCC:553:ASN:N	3:CCC:602:HOH:O	2.27	0.67
1:CCC:553:ASN:ND2	3:CCC:603:HOH:O	2.34	0.60
1:CCC:567:GLN:NE2	3:CCC:601:HOH:O	2.22	0.53

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	CCC	38/38 (100%)	38 (100%)	0	0	100	100
2	DbD	20/31 (64%)	20 (100%)	0	0	100	100
All	All	58/69 (84%)	58 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

### 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	CCC	35/33 (106%)	33 (94%)	2 (6%)	20	1
2	DbD	23/30 (77%)	22 (96%)	1 (4%)	29	3
All	All	58/63 (92%)	55 (95%)	3 (5%)	21	2

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	554	ASN
1	CCC	577	GLN
2	DbD	729	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides 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	CCC	38/38 (100%)	-0.04	1 (2%) 56 58	22, 28, 47, 51	0
2	DbD	22/31 (70%)	0.12	0 100 100	29, 41, 60, 83	0
All	All	60/69 (86%)	0.02	1 (1%) 70 70	22, 30, 59, 83	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	CCC	555	LEU	3.7

### 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 monosaccharides 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.