



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 3, 2024 – 02:27 am GMT

PDB ID : 6H9V  
Title : Crystal structure of deaminated P domain from norovirus strain Saga GII-4 in complex with Fuc  
Authors : Meyer, P.H.O.; Blaum, B.S.  
Deposited on : 2018-08-06  
Resolution : 1.52 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	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

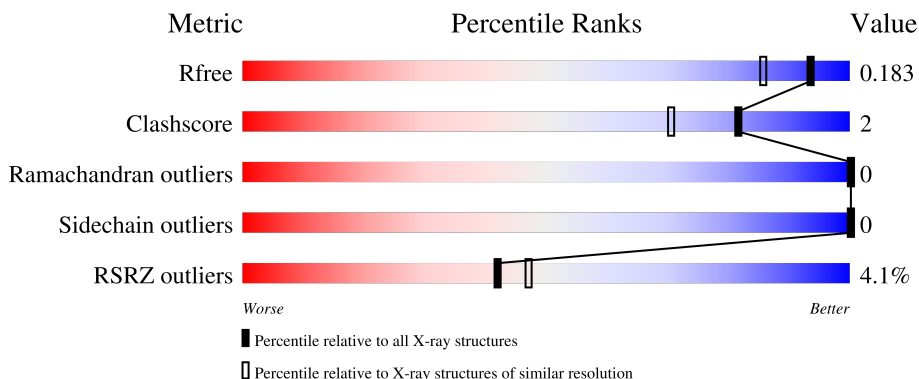
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.52 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	5293 (1.54-1.50)
Clashscore	180529	5759 (1.54-1.50)
Ramachandran outliers	177936	5653 (1.54-1.50)
Sidechain outliers	177891	5650 (1.54-1.50)
RSRZ outliers	164620	5293 (1.54-1.50)

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	309	<div> <div>5%</div> <div> <div></div> <div>94%</div> <div>6%</div> </div> </div>
1	B	309	<div> <div>3%</div> <div> <div></div> <div>95%</div> <div>..</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10493 atoms, of which 4837 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.

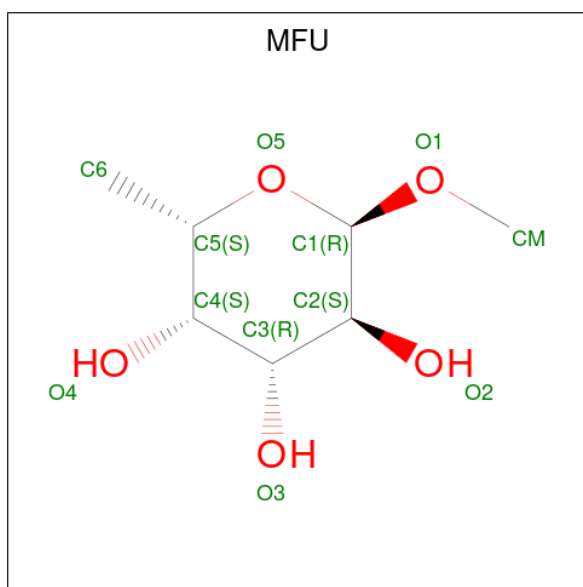
- Molecule 1 is a protein called VP1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	309	Total	C	H	N	O	S	0	32	0
			4987	1625	2441	428	483	10			
1	B	303	Total	C	H	N	O	S	0	19	0
			4836	1569	2368	419	469	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	222	THR	-	expression tag	UNP B5BTR7
A	223	GLY	-	expression tag	UNP B5BTR7
A	224	SER	-	expression tag	UNP B5BTR7
A	373	IAS	ASN	conflict	UNP B5BTR7
B	222	THR	-	expression tag	UNP B5BTR7
B	223	GLY	-	expression tag	UNP B5BTR7
B	224	SER	-	expression tag	UNP B5BTR7
B	373	IAS	ASN	conflict	UNP B5BTR7

- Molecule 2 is methyl alpha-L-fucopyranoside (three-letter code: MFU) (formula: C<sub>7</sub>H<sub>14</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			26	7	14	5		
2	B	1	Total	C	H	O	0	0
			26	7	14	5		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

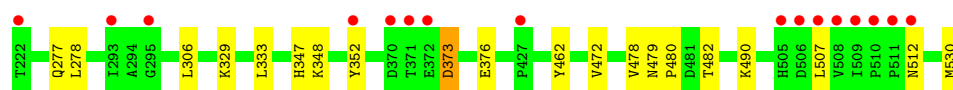
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	295	Total	O	0	0
			295	295		
4	B	321	Total	O	0	0
			321	321		

### 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.

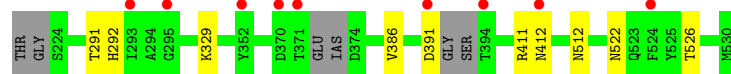
- Molecule 1: VP1

Chain A: 



- Molecule 1: VP1

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.29Å 83.95Å 65.13Å 90.00° 95.78° 90.00°	Depositor
Resolution (Å)	39.58 – 1.52 39.58 – 1.52	Depositor EDS
% Data completeness (in resolution range)	99.0 (39.58-1.52) 99.0 (39.58-1.52)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.20 (at 1.52Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, $R_{free}$	0.147 , 0.183 0.147 , 0.183	Depositor DCC
$R_{free}$ test set	4418 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.5	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 41.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	10493	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.70% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, MFU, IAS

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	0.41	0/2681	0.61	0/3667
1	B	0.43	0/2582	0.60	0/3528
All	All	0.42	0/5263	0.61	0/7195

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

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	2546	2441	2356	18	0
1	B	2468	2368	2329	8	0
2	A	12	14	14	0	0
2	B	12	14	14	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	295	0	0	6	0
4	B	321	0	0	3	0
All	All	5656	4837	4713	24	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 2.

All (24) 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:373:IAS:O	4:A:701:HOH:O	1.96	0.82
1:B:512[A]:ASN:OD1	4:B:701:HOH:O	2.02	0.75
1:A:376[A]:GLU:OE2	4:A:702:HOH:O	2.07	0.71
1:B:526:THR:OG1	4:B:702:HOH:O	2.12	0.67
1:B:329:LYS:NZ	1:B:391:ASP:OD1	2.32	0.59
1:A:306[A]:LEU:HD23	4:A:830:HOH:O	2.02	0.58
1:A:478[A]:VAL:HG11	4:A:745:HOH:O	2.04	0.57
1:A:333[B]:LEU:HD23	1:B:386:VAL:HG11	1.86	0.57
1:B:411:ARG:O	1:B:412:ASN:OD1	2.24	0.55
1:A:278:LEU:HD11	1:A:462:TYR:CE2	2.46	0.51
1:A:479[B]:ASN:ND2	1:A:482:THR:OG1	2.44	0.50
1:A:306[B]:LEU:HD13	4:A:740:HOH:O	2.12	0.50
1:A:329:LYS:HG2	1:A:352[A]:TYR:CD2	2.48	0.49
1:A:333[A]:LEU:HD23	1:A:348:LYS:HA	1.96	0.47
1:A:472[A]:VAL:HG13	1:A:490:LYS:HB3	1.97	0.46
1:A:333[B]:LEU:HD12	1:A:347:HIS:O	2.18	0.44
1:A:507:LEU:HD21	1:A:530:MET:CE	2.48	0.43
1:B:522:ASN:ND2	4:B:707:HOH:O	2.50	0.43
1:A:306[B]:LEU:HG	4:A:926:HOH:O	2.17	0.43
1:A:333[B]:LEU:HD23	1:B:386:VAL:CG1	2.48	0.42
1:A:277:GLN:HG2	1:A:306[B]:LEU:HD11	2.01	0.42
1:A:480:PRO:HG3	1:A:512:ASN:O	2.18	0.42
1:B:291[B]:THR:HG22	1:B:292[B]:HIS:N	2.35	0.41
1:A:507:LEU:HD21	1:A:530:MET:HE3	2.03	0.41

There are no symmetry-related clashes.

## 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	A	337/309 (109%)	326 (97%)	11 (3%)	0	100	100
1	B	317/309 (103%)	313 (99%)	4 (1%)	0	100	100
All	All	654/618 (106%)	639 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	283/267 (106%)	283 (100%)	0	100	100
1	B	278/267 (104%)	278 (100%)	0	100	100
All	All	561/534 (105%)	561 (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. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	IAS	A	373	1	6,7,8	0.57	0	6,8,10	1.63	2 (33%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	IAS	A	373	1	-	0/7/7/8	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	373	IAS	OD1-CG-CB	-2.33	118.64	125.43
1	A	373	IAS	OXT-C-O	-2.23	119.03	124.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	373	IAS	1	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	MFU	A	601	-	12,12,12	1.18	1 (8%)	17,17,17	1.22	2 (11%)
2	MFU	B	601	-	12,12,12	1.33	1 (8%)	17,17,17	1.01	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MFU	A	601	-	-	0/2/22/22	0/1/1/1
2	MFU	B	601	-	-	0/2/22/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	MFU	O5-C1	3.16	1.49	1.41
2	A	601	MFU	O5-C1	2.40	1.48	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	MFU	C6-C5-C4	-2.63	108.21	113.07
2	A	601	MFU	O1-C1-C2	2.44	111.01	108.15

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

### 6.1 Protein, DNA and RNA chains

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	308/309 (99%)	-0.26	16 (5%) 34 37	10, 23, 50, 122	22 (7%)
1	B	303/309 (98%)	-0.34	9 (2%) 52 58	8, 24, 47, 76	16 (5%)
All	All	611/618 (98%)	-0.30	25 (4%) 42 47	8, 24, 49, 122	38 (6%)

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	511	PRO	6.3
1	A	507	LEU	5.1
1	A	512	ASN	5.0
1	A	509	ILE	4.3
1	A	222	THR	4.2
1	A	508	VAL	4.0
1	B	295[A]	GLY	3.6
1	B	371	THR	3.6
1	A	295[A]	GLY	3.5
1	B	524	PHE	3.5
1	A	352[A]	TYR	3.4
1	A	370	ASP	3.2
1	B	293[A]	ILE	3.2
1	B	370	ASP	3.1
1	B	394	THR	3.0
1	A	371	THR	2.9
1	B	352[A]	TYR	2.7
1	A	372	GLU	2.7
1	A	293[A]	ILE	2.4
1	B	412	ASN	2.4
1	A	510	PRO	2.4
1	B	391	ASP	2.4
1	A	506	ASP	2.2
1	A	427	PRO	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	505	HIS	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
1	IAS	A	373	8/9	0.84	0.13	50,66,85,96	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides 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(Å <sup>2</sup> )	Q<0.9
2	MFU	A	601	12/12	0.96	0.07	26,35,50,50	0
2	MFU	B	601	12/12	0.96	0.07	23,32,52,52	0
3	MG	A	602	1/1	0.99	0.04	32,32,32,32	0
3	MG	B	602	1/1	0.99	0.04	32,32,32,32	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.