



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 22, 2024 – 11:33 AM EDT

PDB ID : 5MHR  
Title : T3D reovirus sigma1 complexed with 9BG5 Fab fragments  
Authors : Stehle, T.; Dietrich, M.H.  
Deposited on : 2016-11-25  
Resolution : 3.00 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	:	4.02b-467
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

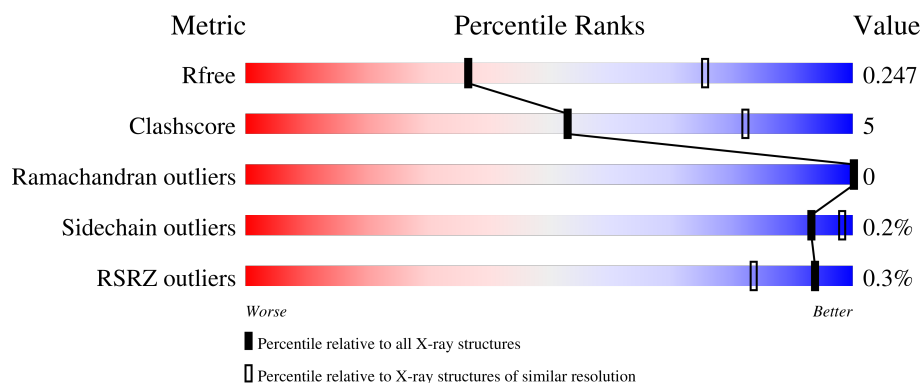
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*





The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	165	
1	B	165	
1	C	165	
1	D	165	
1	E	165	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	165	 87% 12% .
2	G	214	 89% 9% .
2	H	214	 84% 12% .
2	J	214	 42% 8% 50%
2	L	214	 43% 7% 50%
2	O	214	 81% 17% .
2	Q	214	 85% 14% .
3	I	219	 84% 13% .
3	K	219	 51% 5% 44%
3	M	219	 44% 11% 46%
3	N	219	 84% 14% .
3	P	219	 88% 10% .
3	R	219	 88% 9% .

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 22593 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 Viral attachment protein sigma 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	162	Total	C	N	O	S	0	0	0
			1242	795	209	233	5			
1	B	163	Total	C	N	O	S	0	0	0
			1255	801	211	238	5			
1	C	162	Total	C	N	O	S	0	0	0
			1243	796	210	232	5			
1	D	163	Total	C	N	O	S	0	0	0
			1242	795	208	234	5			
1	E	162	Total	C	N	O	S	0	0	0
			1245	797	210	233	5			
1	F	163	Total	C	N	O	S	0	0	0
			1250	797	211	237	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	GLY	-	expression tag	UNP Q86335
A	292	SER	-	expression tag	UNP Q86335
B	291	GLY	-	expression tag	UNP Q86335
B	292	SER	-	expression tag	UNP Q86335
C	291	GLY	-	expression tag	UNP Q86335
C	292	SER	-	expression tag	UNP Q86335
D	291	GLY	-	expression tag	UNP Q86335
D	292	SER	-	expression tag	UNP Q86335
E	291	GLY	-	expression tag	UNP Q86335
E	292	SER	-	expression tag	UNP Q86335
F	291	GLY	-	expression tag	UNP Q86335
F	292	SER	-	expression tag	UNP Q86335

- Molecule 2 is a protein called 9BG5 Fab light chain,LOC100046793 protein,MAb 110 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	211	Total	C	N	O	S	0	0	0
			1522	954	250	311	7			
2	H	205	Total	C	N	O	S	0	0	0
			1455	913	241	293	8			
2	J	108	Total	C	N	O	S	0	0	0
			793	499	130	159	5			
2	L	107	Total	C	N	O	S	0	0	0
			778	490	124	159	5			
2	O	210	Total	C	N	O	S	0	0	0
			1470	921	245	296	8			
2	Q	210	Total	C	N	O	S	0	0	0
			1422	885	240	289	8			

- Molecule 3 is a protein called 9BG5 Fab heavy chain.

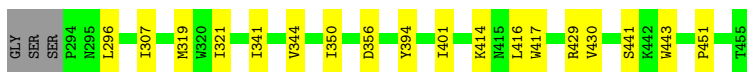
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	213	Total	C	N	O	S	0	0	0
			1491	953	240	291	7			
3	K	122	Total	C	N	O	S	0	0	0
			887	569	137	177	4			
3	M	119	Total	C	N	O	S	0	0	0
			853	550	132	168	3			
3	N	214	Total	C	N	O	S	0	0	0
			1544	991	247	300	6			
3	P	213	Total	C	N	O	S	0	0	0
			1474	938	237	293	6			
3	R	213	Total	C	N	O	S	0	0	0
			1427	909	234	278	6			

### 3 Residue-property plots


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: Viral attachment protein sigma 1

Chain A: 




- Molecule 1: Viral attachment protein sigma 1

Chain B: 




- Molecule 1: Viral attachment protein sigma 1

Chain C: 




- Molecule 1: Viral attachment protein sigma 1

Chain D: 




- Molecule 1: Viral attachment protein sigma 1

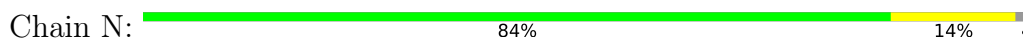
Chain E: 



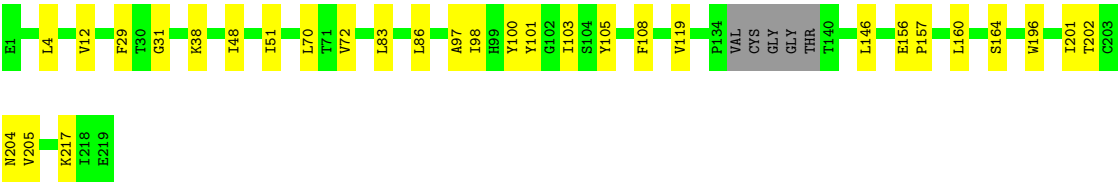
- Molecule 1: Viral attachment protein sigma 1

Chain F: 

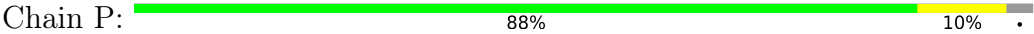




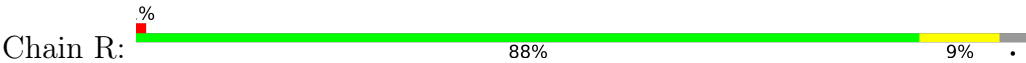




● Molecule 3: 9BG5 Fab heavy chain



● Molecule 3: 9BG5 Fab heavy chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.16Å 109.31Å 131.70Å 103.13° 113.56° 103.48°	Depositor
Resolution (Å)	49.05 – 3.00 49.05 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.8 (49.05-3.00) 98.8 (49.05-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.89 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.212 , 0.247 0.212 , 0.247	Depositor DCC
$R_{free}$ test set	4734 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.9	Xtriage
Anisotropy	0.145	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 31.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	22593	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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

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	0.29	0/1278	0.50	0/1744
1	B	0.29	0/1291	0.51	0/1761
1	C	0.30	0/1279	0.52	0/1746
1	D	0.28	0/1278	0.51	0/1745
1	E	0.29	0/1281	0.51	0/1747
1	F	0.30	0/1286	0.50	0/1756
2	G	0.26	0/1557	0.47	0/2129
2	H	0.28	0/1486	0.46	0/2025
2	J	0.29	0/809	0.52	0/1098
2	L	0.26	0/794	0.48	0/1080
2	O	0.28	0/1501	0.48	0/2055
2	Q	0.26	0/1450	0.46	0/1985
3	I	0.28	0/1535	0.49	0/2116
3	K	0.28	0/912	0.47	0/1253
3	M	0.27	0/879	0.49	0/1212
3	N	0.28	0/1589	0.51	0/2185
3	P	0.27	0/1518	0.45	0/2099
3	R	0.26	0/1469	0.46	0/2032
All	All	0.28	0/23192	0.49	0/31768

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	1242	0	1189	11	0
1	B	1255	0	1204	14	0
1	C	1243	0	1192	17	0
1	D	1242	0	1181	13	0
1	E	1245	0	1195	14	0
1	F	1250	0	1190	12	0
2	G	1522	0	1379	12	0
2	H	1455	0	1279	14	0
2	J	793	0	771	12	0
2	L	778	0	748	10	0
2	O	1470	0	1305	27	0
2	Q	1422	0	1224	17	0
3	I	1491	0	1273	19	0
3	K	887	0	766	8	0
3	M	853	0	716	16	0
3	N	1544	0	1380	18	0
3	P	1474	0	1229	15	0
3	R	1427	0	1124	12	0
All	All	22593	0	20345	232	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 5.

The worst 5 of 232 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)
2:G:108:ARG:NH2	2:G:170:ASP:OD1	2.25	0.70
3:N:83:LEU:HB3	3:N:86:LEU:HD21	1.72	0.69
2:J:25:ALA:O	2:J:70:THR:OG1	2.11	0.68
2:Q:31:SER:O	2:Q:52:THR:OG1	2.10	0.68
3:I:11:LEU:HD22	3:I:155:PRO:HG3	1.76	0.68

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	160/165 (97%)	157 (98%)	3 (2%)	0	100	100
1	B	161/165 (98%)	158 (98%)	3 (2%)	0	100	100
1	C	160/165 (97%)	156 (98%)	4 (2%)	0	100	100
1	D	161/165 (98%)	157 (98%)	4 (2%)	0	100	100
1	E	160/165 (97%)	157 (98%)	3 (2%)	0	100	100
1	F	161/165 (98%)	157 (98%)	4 (2%)	0	100	100
2	G	209/214 (98%)	206 (99%)	3 (1%)	0	100	100
2	H	197/214 (92%)	193 (98%)	4 (2%)	0	100	100
2	J	106/214 (50%)	103 (97%)	3 (3%)	0	100	100
2	L	105/214 (49%)	102 (97%)	3 (3%)	0	100	100
2	O	206/214 (96%)	201 (98%)	5 (2%)	0	100	100
2	Q	208/214 (97%)	202 (97%)	6 (3%)	0	100	100
3	I	209/219 (95%)	204 (98%)	5 (2%)	0	100	100
3	K	118/219 (54%)	116 (98%)	2 (2%)	0	100	100
3	M	117/219 (53%)	113 (97%)	4 (3%)	0	100	100
3	N	210/219 (96%)	208 (99%)	2 (1%)	0	100	100
3	P	209/219 (95%)	204 (98%)	5 (2%)	0	100	100
3	R	207/219 (94%)	202 (98%)	5 (2%)	0	100	100
All	All	3064/3588 (85%)	2996 (98%)	68 (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	134/138 (97%)	133 (99%)	1 (1%)	84	94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	137/138 (99%)	137 (100%)	0	100	100
1	C	134/138 (97%)	133 (99%)	1 (1%)	84	94
1	D	133/138 (96%)	133 (100%)	0	100	100
1	E	134/138 (97%)	134 (100%)	0	100	100
1	F	135/138 (98%)	135 (100%)	0	100	100
2	G	156/187 (83%)	156 (100%)	0	100	100
2	H	140/187 (75%)	139 (99%)	1 (1%)	84	94
2	J	85/187 (46%)	84 (99%)	1 (1%)	71	90
2	L	84/187 (45%)	84 (100%)	0	100	100
2	O	143/187 (76%)	142 (99%)	1 (1%)	84	94
2	Q	131/187 (70%)	131 (100%)	0	100	100
3	I	136/188 (72%)	136 (100%)	0	100	100
3	K	85/188 (45%)	85 (100%)	0	100	100
3	M	76/188 (40%)	76 (100%)	0	100	100
3	N	151/188 (80%)	151 (100%)	0	100	100
3	P	135/188 (72%)	135 (100%)	0	100	100
3	R	114/188 (61%)	114 (100%)	0	100	100
All	All	2243/3078 (73%)	2238 (100%)	5 (0%)	93	98

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	356	ASP
1	C	356	ASP
2	H	91	GLN
2	J	91	GLN
2	O	91	GLN

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 [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

### 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	162/165 (98%)	-0.71	0 <span>100</span> <span>100</span>	29, 43, 81, 116	0
1	B	163/165 (98%)	-0.71	1 (0%) <span>89</span> <span>72</span>	29, 44, 82, 117	0
1	C	162/165 (98%)	-0.76	0 <span>100</span> <span>100</span>	26, 43, 78, 130	0
1	D	163/165 (98%)	-0.74	0 <span>100</span> <span>100</span>	29, 46, 92, 134	0
1	E	162/165 (98%)	-0.70	0 <span>100</span> <span>100</span>	32, 48, 97, 136	0
1	F	163/165 (98%)	-0.73	0 <span>100</span> <span>100</span>	35, 56, 98, 134	0
2	G	211/214 (98%)	-0.53	1 (0%) <span>91</span> <span>75</span>	40, 75, 131, 156	0
2	H	205/214 (95%)	-0.52	0 <span>100</span> <span>100</span>	35, 71, 141, 170	0
2	J	108/214 (50%)	-0.75	0 <span>100</span> <span>100</span>	36, 49, 77, 101	0
2	L	107/214 (50%)	-0.77	0 <span>100</span> <span>100</span>	36, 49, 73, 81	0
2	O	210/214 (98%)	-0.52	0 <span>100</span> <span>100</span>	43, 77, 154, 197	0
2	Q	210/214 (98%)	-0.47	0 <span>100</span> <span>100</span>	34, 79, 163, 181	0
3	I	213/219 (97%)	-0.40	3 (1%) <span>75</span> <span>49</span>	32, 73, 142, 175	0
3	K	122/219 (55%)	-0.74	0 <span>100</span> <span>100</span>	41, 78, 115, 128	0
3	M	119/219 (54%)	-0.76	1 (0%) <span>86</span> <span>65</span>	45, 84, 114, 139	0
3	N	214/219 (97%)	-0.64	0 <span>100</span> <span>100</span>	39, 74, 120, 145	0
3	P	213/219 (97%)	-0.50	0 <span>100</span> <span>100</span>	51, 107, 142, 167	0
3	R	213/219 (97%)	-0.41	2 (0%) <span>84</span> <span>63</span>	40, 101, 166, 182	0
All	All	3120/3588 (86%)	-0.61	8 (0%) <span>94</span> <span>84</span>	26, 63, 140, 197	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	150	VAL	3.4
3	R	206	ALA	3.3
3	I	214	VAL	2.4

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	I	196	TRP	2.3
3	M	74	LYS	2.2

## 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.