



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 11:07 pm BST

PDB ID : 3G3A  
Title : Structure of a lamprey variable lymphocyte receptor in complex with a protein antigen  
Authors : Deng, L.; Velikovsky, C.A.; Mariuzza, R.A.  
Deposited on : 2009-02-02  
Resolution : 2.20 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 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	:	2.11
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)	:	2.11

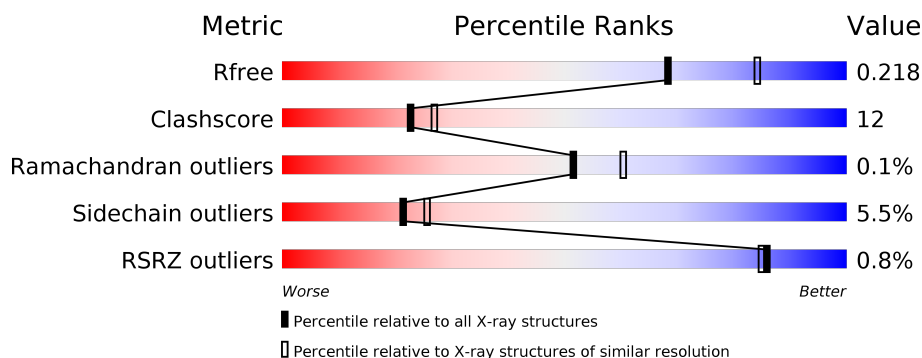
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 on 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	178	
1	C	178	
1	E	178	
1	G	178	
2	B	129	
2	D	129	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	129	 75% 19% . .
2	H	129	 2% 67% 28% . .

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9537 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 Variable lymphocyte receptor VLRB.2D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	167	Total	C	N	O	S	0	0	0
			1274	797	217	252	8			
1	C	167	Total	C	N	O	S	0	0	0
			1274	797	217	252	8			
1	E	167	Total	C	N	O	S	0	0	0
			1274	797	217	252	8			
1	G	165	Total	C	N	O	S	0	0	0
			1259	789	214	248	8			

- Molecule 2 is a protein called Lysozyme C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	126	Total	C	N	O	S	0	0	0
			973	597	186	180	10			
2	D	126	Total	C	N	O	S	0	0	0
			973	597	186	180	10			
2	F	126	Total	C	N	O	S	0	0	0
			971	596	186	179	10			
2	H	124	Total	C	N	O	S	0	0	0
			961	591	184	176	10			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	93	Total	O	0	0
			93	93		
3	B	78	Total	O	0	0
			78	78		
3	C	81	Total	O	0	0
			81	81		
3	D	71	Total	O	0	0
			71	71		

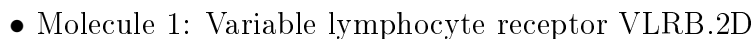
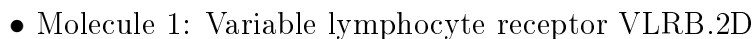
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	62	Total 62	O 62	0	0
3	F	59	Total 59	O 59	0	0
3	G	67	Total 67	O 67	0	0
3	H	67	Total 67	O 67	0	0

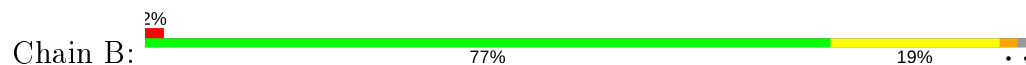


- Molecule 1: Variable lymphocyte receptor VLRB.2D

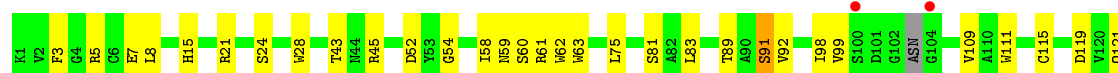




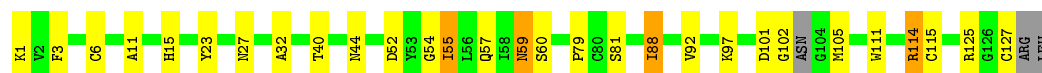
• Molecule 2: Lysozyme C



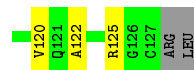
• Molecule 2: Lysozyme C



• Molecule 2: Lysozyme C



• Molecule 2: Lysozyme C



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.30Å 133.65Å 90.44Å 90.00° 114.69° 90.00°	Depositor
Resolution (Å)	50.00 – 2.20 30.95 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.4 (50.00-2.20) 98.5 (30.95-2.20)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.09 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.220 , 0.271 0.224 , 0.218	Depositor DCC
$R_{free}$ test set	4437 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.4	Xtriage
Anisotropy	0.490	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 26.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.094 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	9537	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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	0.87	0/1300	0.85	1/1779 (0.1%)
1	C	0.83	0/1300	0.87	0/1779
1	E	0.83	0/1300	0.88	1/1779 (0.1%)
1	G	0.79	0/1284	0.83	0/1756
2	B	1.00	0/992	0.89	1/1340 (0.1%)
2	D	0.95	0/992	0.86	2/1340 (0.1%)
2	F	0.91	0/990	0.87	0/1337
2	H	0.95	0/980	0.87	1/1324 (0.1%)
All	All	0.89	0/9138	0.86	6/12434 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	14	ASP	CB-CG-OD1	5.76	123.48	118.30
2	B	45	ARG	NE-CZ-NH2	-5.74	117.43	120.30
1	E	55	LEU	CA-CB-CG	5.66	128.33	115.30
2	D	45	ARG	NE-CZ-NH1	5.47	123.03	120.30
2	H	114	ARG	NE-CZ-NH2	-5.46	117.57	120.30

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	A	1274	0	1259	26	0
1	C	1274	0	1257	25	0
1	E	1274	0	1257	34	0
1	G	1259	0	1243	18	0
2	B	973	0	930	26	0
2	D	973	0	930	24	0
2	F	971	0	926	25	0
2	H	961	0	921	31	0
3	A	93	0	0	5	0
3	B	78	0	0	6	0
3	C	81	0	0	5	0
3	D	71	0	0	6	0
3	E	62	0	0	8	0
3	F	59	0	0	4	0
3	G	67	0	0	1	0
3	H	67	0	0	2	0
All	All	9537	0	8723	203	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 12.

The worst 5 of 203 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:G:51:ARG:HG2	3:G:401:HOH:O	1.51	1.10
2:F:27:ASN:HB3	2:F:105:MET:HE1	1.34	1.09
2:H:27:ASN:HB3	2:H:105:MET:HE1	1.45	0.97
2:F:27:ASN:HB3	2:F:105:MET:CE	1.96	0.95
1:G:69:PRO:HG2	1:G:72:VAL:HG13	1.51	0.91

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	165/178 (93%)	152 (92%)	13 (8%)	0	100	100
1	C	165/178 (93%)	150 (91%)	14 (8%)	1 (1%)	25	26
1	E	165/178 (93%)	147 (89%)	18 (11%)	0	100	100
1	G	161/178 (90%)	151 (94%)	10 (6%)	0	100	100
2	B	122/129 (95%)	117 (96%)	5 (4%)	0	100	100
2	D	122/129 (95%)	119 (98%)	3 (2%)	0	100	100
2	F	122/129 (95%)	119 (98%)	3 (2%)	0	100	100
2	H	120/129 (93%)	114 (95%)	6 (5%)	0	100	100
All	All	1142/1228 (93%)	1069 (94%)	72 (6%)	1 (0%)	51	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	160	ALA

### 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	150/157 (96%)	143 (95%)	7 (5%)	26	33
1	C	150/157 (96%)	145 (97%)	5 (3%)	38	49
1	E	150/157 (96%)	138 (92%)	12 (8%)	12	12
1	G	148/157 (94%)	137 (93%)	11 (7%)	13	14
2	B	102/105 (97%)	97 (95%)	5 (5%)	25	31
2	D	102/105 (97%)	98 (96%)	4 (4%)	32	41
2	F	101/105 (96%)	95 (94%)	6 (6%)	19	23
2	H	101/105 (96%)	96 (95%)	5 (5%)	24	30
All	All	1004/1048 (96%)	949 (94%)	55 (6%)	21	26

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	45	GLU
1	E	94	ARG
2	H	24	SER
1	E	52	LEU
1	E	68	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	93	ASN
1	E	54	GLN
2	H	27	ASN
2	D	113	ASN
1	E	62	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 carbohydrates 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	167/178 (93%)	-0.22	0 100 100	12, 24, 38, 52	0
1	C	167/178 (93%)	-0.08	0 100 100	13, 27, 44, 54	0
1	E	167/178 (93%)	0.16	3 (1%) 68 66	18, 32, 49, 57	1 (0%)
1	G	165/178 (92%)	-0.23	0 100 100	14, 25, 37, 46	1 (0%)
2	B	126/129 (97%)	-0.18	2 (1%) 72 70	10, 20, 35, 43	4 (3%)
2	D	126/129 (97%)	-0.13	2 (1%) 72 70	12, 23, 36, 42	3 (2%)
2	F	126/129 (97%)	-0.11	0 100 100	14, 24, 40, 49	0
2	H	124/129 (96%)	-0.13	2 (1%) 72 70	13, 22, 34, 41	2 (1%)
All	All	1168/1228 (95%)	-0.11	9 (0%) 86 85	10, 25, 40, 57	11 (0%)

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	104	GLY	4.7
2	H	99	VAL	4.2
2	B	104	GLY	4.0
1	E	1	ALA	3.7
2	B	99	VAL	3.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates 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.