



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2025 – 09:08 am GMT

PDB ID : 9H4R  
Title : Structure of fertilization-blocking monoclonal antibody IE-3 VHVL bound to the ZP-N1 domain of mouse ZP2 (crystal form I)  
Authors : Dioguardi, E.; De Sanctis, D.; Jovine, L.  
Deposited on : 2024-10-21  
Resolution : 1.53 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	:	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)	:	2.41

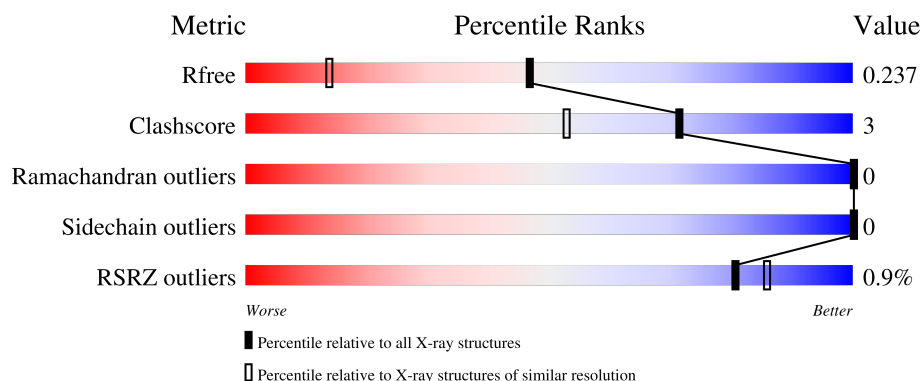
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.53 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	3511 (1.56-1.52)
Clashscore	180529	3784 (1.56-1.52)
Ramachandran outliers	177936	3720 (1.56-1.52)
Sidechain outliers	177891	3717 (1.56-1.52)
RSRZ outliers	164620	3510 (1.56-1.52)

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	112	 79% 7% 14%
1	B	112	 82% 14%
2	H	123	 90% 7% .
2	X	123	 96% . .
3	L	117	 91% 8% .

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	Y	117	 A horizontal bar chart showing the quality of chain Y. The bar is divided into two segments: a green segment representing 84% and a yellow segment representing 15%. A small grey dot is at the end of the bar. The text '84%' is centered under the green segment and '15%' is centered under the yellow segment.

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5768 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 Zona pellucida sperm-binding protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	96	Total	C	N	O	S	0	0	0
			776	497	123	150	6			
1	B	96	Total	C	N	O	S	0	0	0
			775	498	123	148	6			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	83	SER	ASN	engineered mutation	UNP P20239
A	139	LEU	-	expression tag	UNP P20239
A	140	GLU	-	expression tag	UNP P20239
A	141	HIS	-	expression tag	UNP P20239
A	142	HIS	-	expression tag	UNP P20239
A	143	HIS	-	expression tag	UNP P20239
A	144	HIS	-	expression tag	UNP P20239
A	145	HIS	-	expression tag	UNP P20239
A	146	HIS	-	expression tag	UNP P20239
B	83	SER	ASN	engineered mutation	UNP P20239
B	139	LEU	-	expression tag	UNP P20239
B	140	GLU	-	expression tag	UNP P20239
B	141	HIS	-	expression tag	UNP P20239
B	142	HIS	-	expression tag	UNP P20239
B	143	HIS	-	expression tag	UNP P20239
B	144	HIS	-	expression tag	UNP P20239
B	145	HIS	-	expression tag	UNP P20239
B	146	HIS	-	expression tag	UNP P20239

- Molecule 2 is a protein called Heavy chain variable (VH) domain of anti-ZP2 monoclonal antibody IE-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	120	Total	C	N	O	S	0	0	0
			912	568	151	188	5			
2	X	120	Total	C	N	O	S	0	0	0
			912	568	151	188	5			

- Molecule 3 is a protein called Light chain variable (VL) domain of anti-ZP2 monoclonal antibody IE-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	115	Total	C	N	O	S	0	0	0
			883	553	147	180	3			
3	Y	116	Total	C	N	O	S	0	0	0
			894	560	152	179	3			


- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	132	Total	O	0	0
			132	132		
4	H	82	Total	O	0	0
			82	82		
4	L	77	Total	O	0	0
			77	77		
4	B	130	Total	O	0	0
			130	130		
4	X	100	Total	O	0	0
			100	100		
4	Y	95	Total	O	0	0
			95	95		

### 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.

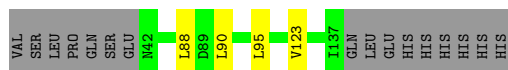
- Molecule 1: Zona pellucida sperm-binding protein 2

Chain A: 




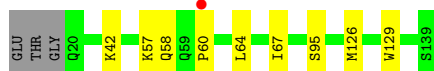
- Molecule 1: Zona pellucida sperm-binding protein 2

Chain B: 



- Molecule 2: Heavy chain variable (VH) domain of anti-ZP2 monoclonal antibody IE-3

Chain H: 

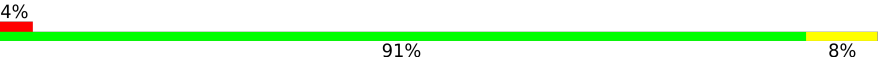


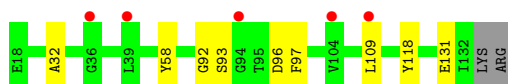
- Molecule 2: Heavy chain variable (VH) domain of anti-ZP2 monoclonal antibody IE-3

Chain X: 



- Molecule 3: Light chain variable (VL) domain of anti-ZP2 monoclonal antibody IE-3

Chain L: 



- Molecule 3: Light chain variable (VL) domain of anti-ZP2 monoclonal antibody IE-3

Chain Y: 

84%

15%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.64Å 92.29Å 91.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.16 – 1.53 29.16 – 1.53	Depositor EDS
% Data completeness (in resolution range)	93.4 (29.16-1.53) 91.5 (29.16-1.53)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.22 (at 1.53Å)	Xtriage
Refinement program	PHENIX 1.21.2_5419+SVN	Depositor
R, $R_{free}$	0.206 , 0.233 0.210 , 0.237	Depositor DCC
$R_{free}$ test set	6015 reflections (5.51%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.7	Xtriage
Anisotropy	0.271	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 37.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.258 for k,h,-l 0.044 for -h,-l,-k 0.046 for l,-k,h 0.044 for k,l,h 0.044 for l,h,k	Xtriage
Reported twinning fraction	0.300 for k,h,-l	Depositor
Outliers	0 of 115130 reflections	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	5768	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.22% 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	0.26	0/795	0.47	0/1079
1	B	0.26	0/794	0.47	0/1078
2	H	0.26	0/932	0.45	0/1262
2	X	0.26	0/932	0.45	0/1262
3	L	0.26	0/902	0.47	0/1226
3	Y	0.25	0/913	0.48	0/1239
All	All	0.26	0/5268	0.46	0/7146

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	A	776	0	740	6	0
1	B	775	0	745	4	0
2	H	912	0	858	5	0
2	X	912	0	858	2	0
3	L	883	0	847	6	0
3	Y	894	0	867	12	0
4	A	132	0	0	2	0
4	B	130	0	0	0	0
4	H	82	0	0	1	0
4	L	77	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	X	100	0	0	0	0
4	Y	95	0	0	3	0
All	All	5768	0	4915	32	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 3.

The worst 5 of 32 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)
3:Y:19:THR:HG22	4:Y:277:HOH:O	1.73	0.87
3:Y:68:GLN:NE2	4:Y:204:HOH:O	2.27	0.66
1:A:88:LEU:HB3	1:B:88:LEU:HB3	1.78	0.65
2:H:42:LYS:NZ	2:H:95:SER:O	2.38	0.55
3:Y:63:GLN:HB2	3:Y:73:LEU:HD11	1.92	0.52

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	94/112 (84%)	93 (99%)	1 (1%)	0	100	100
1	B	94/112 (84%)	93 (99%)	1 (1%)	0	100	100
2	H	118/123 (96%)	118 (100%)	0	0	100	100
2	X	118/123 (96%)	117 (99%)	1 (1%)	0	100	100
3	L	113/117 (97%)	108 (96%)	5 (4%)	0	100	100
3	Y	114/117 (97%)	110 (96%)	4 (4%)	0	100	100
All	All	651/704 (92%)	639 (98%)	12 (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	88/104 (85%)	88 (100%)	0	100	100
1	B	88/104 (85%)	88 (100%)	0	100	100
2	H	100/102 (98%)	100 (100%)	0	100	100
2	X	100/102 (98%)	100 (100%)	0	100	100
3	L	100/102 (98%)	100 (100%)	0	100	100
3	Y	101/102 (99%)	101 (100%)	0	100	100
All	All	577/616 (94%)	577 (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 ⓘ

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

### 5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

### 5.6 Ligand geometry ⓘ

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	A	96/112 (85%)	-0.91	0 <span>100</span> <span>100</span>	25, 28, 37, 79	0
1	B	96/112 (85%)	-0.91	0 <span>100</span> <span>100</span>	24, 29, 44, 78	0
2	H	120/123 (97%)	-0.50	1 (0%) <span>82</span> <span>87</span>	28, 37, 58, 68	0
2	X	120/123 (97%)	-0.79	0 <span>100</span> <span>100</span>	25, 32, 46, 68	0
3	L	115/117 (98%)	-0.02	5 (4%) <span>40</span> <span>49</span>	27, 46, 89, 117	0
3	Y	116/117 (99%)	-0.56	0 <span>100</span> <span>100</span>	25, 36, 51, 88	0
All	All	663/704 (94%)	-0.60	6 (0%) <span>81</span> <span>86</span>	24, 34, 64, 117	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	36	GLY	3.0
2	H	60	PRO	2.3
3	L	109	LEU	2.2
3	L	94	GLY	2.1
3	L	104	VAL	2.0

### 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.