



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 03:04 PM UTC

PDB ID : 7SO7 / pdb\_00007so7  
Title : Novel structural insights for a pair of monoclonal antibodies recognizing non-overlapping epitopes of the glucosyltransferase domain of Clostridium difficile toxin B  
Authors : Liu, J.  
Deposited on : 2021-10-29  
Resolution : 3.59 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 : 4-5-2 with Phenix2.0  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

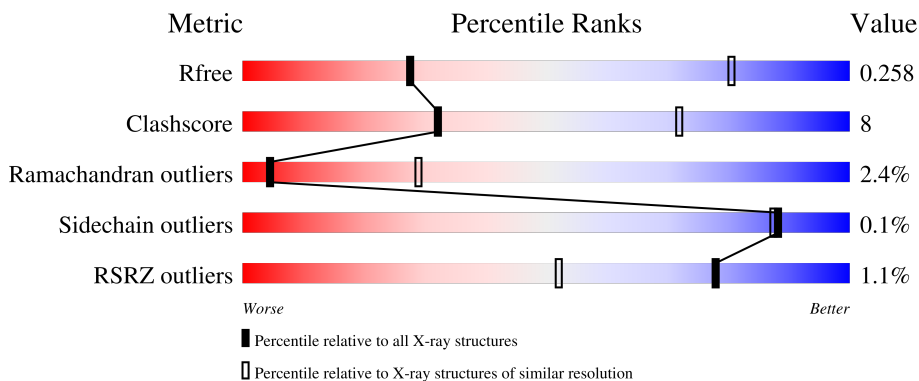
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.59 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))
$R_{free}$	180053	1747 (3.70-3.50)
Clashscore	190562	1827 (3.70-3.50)
Ramachandran outliers	187476	1773 (3.70-3.50)
Sidechain outliers	187428	1772 (3.70-3.50)
RSRZ outliers	180081	1745 (3.70-3.50)

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	538	 2% 79% 18% ..
1	B	538	 % 83% 15% .
2	H	213	 77% 21% .
2	X	213	 75% 23% .
3	L	218	 77% 22% .

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	Y	218	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into two segments: a green segment on the left labeled '75%' and a yellow segment on the right labeled '23%'. A small red square is at the beginning of the bar, and a small black dot is at the end. A '%' symbol is positioned above the start of the bar.</p>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15130 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 Toxin B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	528	Total	C	N	O	S	0	0	0
			4315	2737	696	865	17			
1	B	529	Total	C	N	O	S	0	0	0
			4326	2743	701	865	17			

- Molecule 2 is a protein called Fab B1 HC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	X	213	Total	C	N	O	S	0	0	0
			1603	1012	270	313	8			
2	H	213	Total	C	N	O	S	0	0	0
			1603	1012	270	313	8			

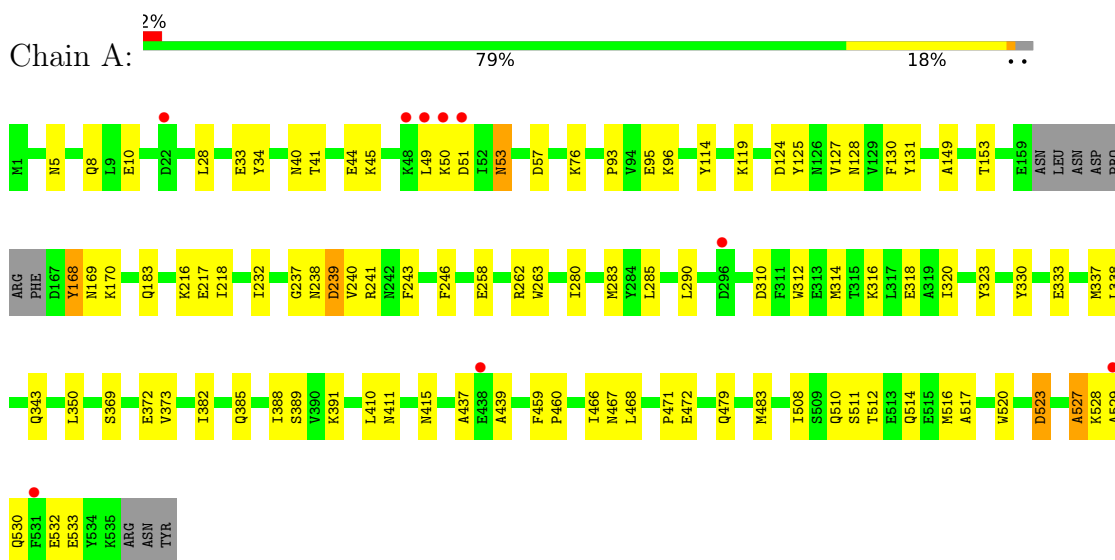
- Molecule 3 is a protein called FAB B1 LC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	Y	218	Total	C	N	O	S	0	0	0
			1642	1037	278	323	4			
3	L	218	Total	C	N	O	S	0	0	0
			1641	1036	278	323	4			

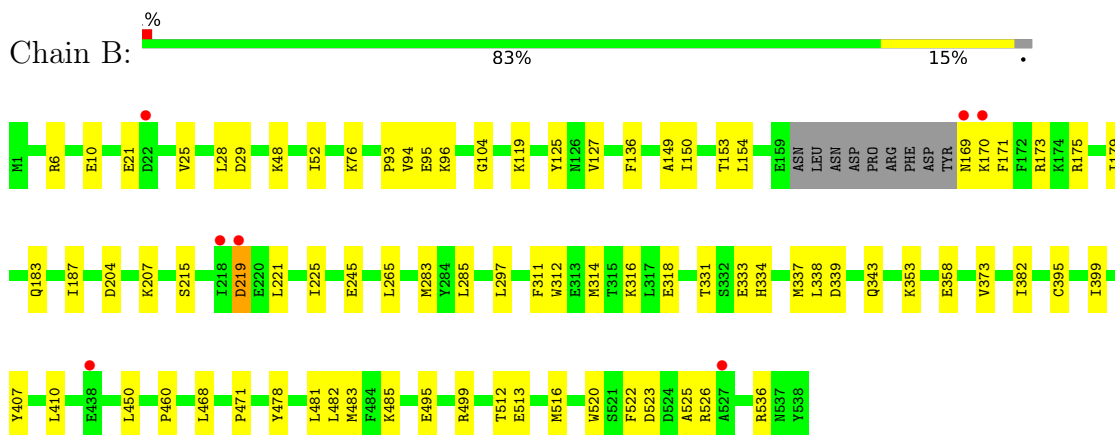
### 3 Residue-property plots [i](#)

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.

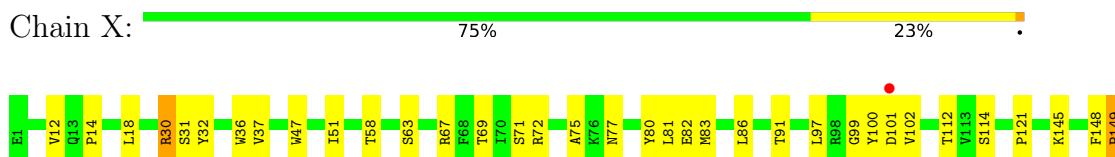
- Molecule 1: Toxin B



- Molecule 1: Toxin B

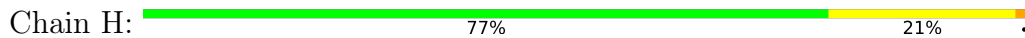


- Molecule 2: Fab B1 HC

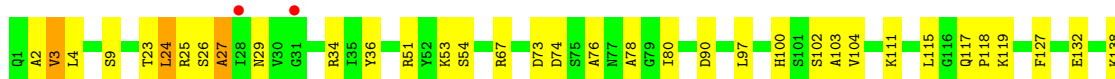
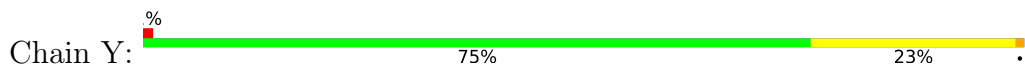




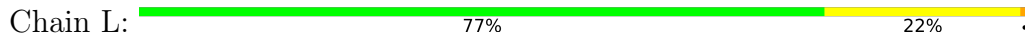
• Molecule 2: Fab B1 HC



• Molecule 3: FAB B1 LC



• Molecule 3: FAB B1 LC



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	210.54Å 320.08Å 65.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	74.80 – 3.59 74.80 – 3.59	Depositor EDS
% Data completeness (in resolution range)	99.9 (74.80-3.59) 99.9 (74.80-3.59)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.06 (at 3.58Å)	Xtrriage
Refinement program	PHENIX 1.12_2829	Depositor
R, $R_{free}$	0.209 , 0.258 0.213 , 0.258	Depositor DCC
$R_{free}$ test set	2000 reflections (3.76%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	91.4	Xtrriage
Anisotropy	0.410	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 71.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	15130	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.09% 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	0.14	0/4391	0.36	0/5923
1	B	0.14	0/4402	0.34	0/5937
2	H	0.15	0/1641	0.41	0/2234
2	X	0.17	0/1641	0.44	0/2234
3	L	0.19	0/1683	0.44	0/2293
3	Y	0.16	0/1684	0.43	0/2295
All	All	0.15	0/15442	0.38	0/20916

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	H	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	148	PHE	Peptide

### 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	4315	0	4208	66	0
1	B	4326	0	4223	48	0
2	H	1603	0	1568	35	0
2	X	1603	0	1568	37	0
3	L	1641	0	1611	35	0
3	Y	1642	0	1615	35	0
All	All	15130	0	14793	238	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 8.

The worst 5 of 238 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:H:211:LYS:NZ	3:L:132:GLU:OE2	1.81	1.12
3:L:34:ARG:HG2	3:L:53:LYS:HB2	1.55	0.88
2:X:30:ARG:O	2:X:32:TYR:N	2.17	0.78
1:B:95:GLU:HG3	3:Y:34:ARG:HH22	1.50	0.77
3:L:73:ASP:HB2	3:L:80:ILE:HD11	1.67	0.76

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	524/538 (97%)	471 (90%)	43 (8%)	10 (2%)	6	33
1	B	525/538 (98%)	478 (91%)	44 (8%)	3 (1%)	21	54
2	H	209/213 (98%)	187 (90%)	14 (7%)	8 (4%)	2	21
2	X	209/213 (98%)	190 (91%)	11 (5%)	8 (4%)	2	21
3	L	216/218 (99%)	181 (84%)	28 (13%)	7 (3%)	3	24

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	Y	216/218 (99%)	174 (81%)	33 (15%)	9 (4%)	2	19
All	All	1899/1938 (98%)	1681 (88%)	173 (9%)	45 (2%)	4	29

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	168	TYR
1	A	217	GLU
1	A	240	VAL
1	A	241	ARG
2	X	30	ARG

### 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	A	483/493 (98%)	483 (100%)	0	100	100
1	B	484/493 (98%)	484 (100%)	0	100	100
2	H	180/180 (100%)	180 (100%)	0	100	100
2	X	180/180 (100%)	180 (100%)	0	100	100
3	L	182/182 (100%)	181 (100%)	1 (0%)	81	80
3	Y	182/182 (100%)	182 (100%)	0	100	100
All	All	1691/1710 (99%)	1690 (100%)	1 (0%)	88	87

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	171	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	39	GLN
3	L	60	GLN
3	L	203	GLN
3	L	176	GLN
3	Y	29	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 oligosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	X	1
2	H	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	X	128:PRO	C	135:GLY	N	9.41
1	H	128:PRO	C	135:GLY	N	8.56

## 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	528/538 (98%)	-0.00	9 (1%) 69 41	53, 91, 151, 208	0
1	B	529/538 (98%)	-0.09	7 (1%) 75 48	51, 89, 153, 211	0
2	H	213/213 (100%)	-0.05	1 (0%) 87 66	55, 88, 126, 180	0
2	X	213/213 (100%)	-0.13	1 (0%) 87 66	54, 87, 132, 185	0
3	L	218/218 (100%)	-0.04	1 (0%) 87 66	58, 85, 132, 225	0
3	Y	218/218 (100%)	0.00	2 (0%) 81 56	56, 84, 128, 195	0
All	All	1919/1938 (99%)	-0.05	21 (1%) 78 52	51, 88, 145, 225	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	51	ASP	4.1
1	A	48	LYS	4.0
3	Y	28	ILE	3.7
1	B	527	ALA	3.2
3	L	28	ILE	2.9

### 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 oligosaccharides 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.