



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2023 – 05:06 PM EDT

PDB ID : 2A97  
Title : Crystal structure of catalytic domain of Clostridium botulinum neurotoxin serotype F  
Authors : Agarwal, R.; Binz, T.; Swaminathan, S.  
Deposited on : 2005-07-11  
Resolution : 1.80 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 : 2.35  
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) : 2.35

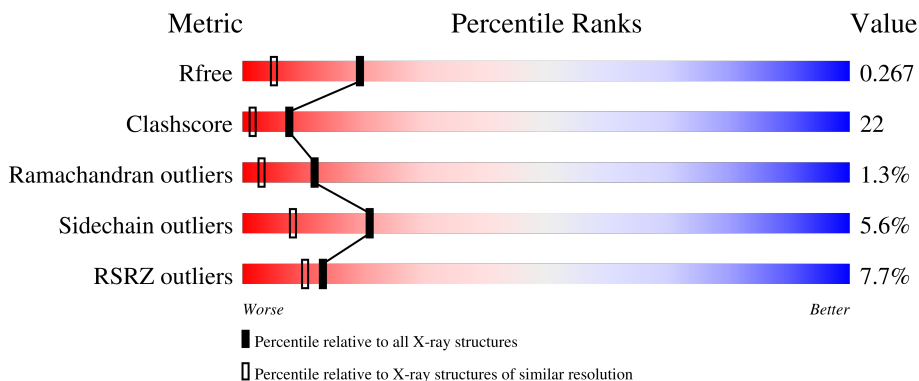
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	439	
1	B	439	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6586 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 Botulinum neurotoxin type F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	392	3168	2038	505	618	7	0	0	0
1	B	392	3168	2038	505	618	7	0	0	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0

- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	5	Total 5	Cd 5	0	0
3	B	2	Total 2	Cd 2	0	0

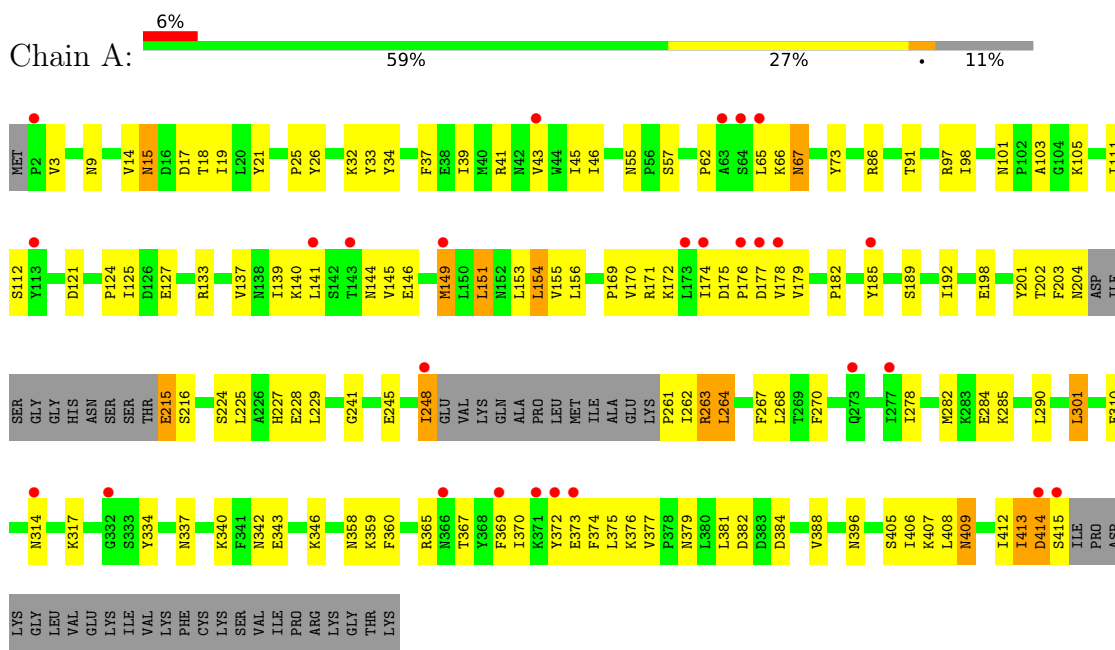
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	109	Total 109	O 109	0	0
4	B	132	Total 132	O 132	0	0

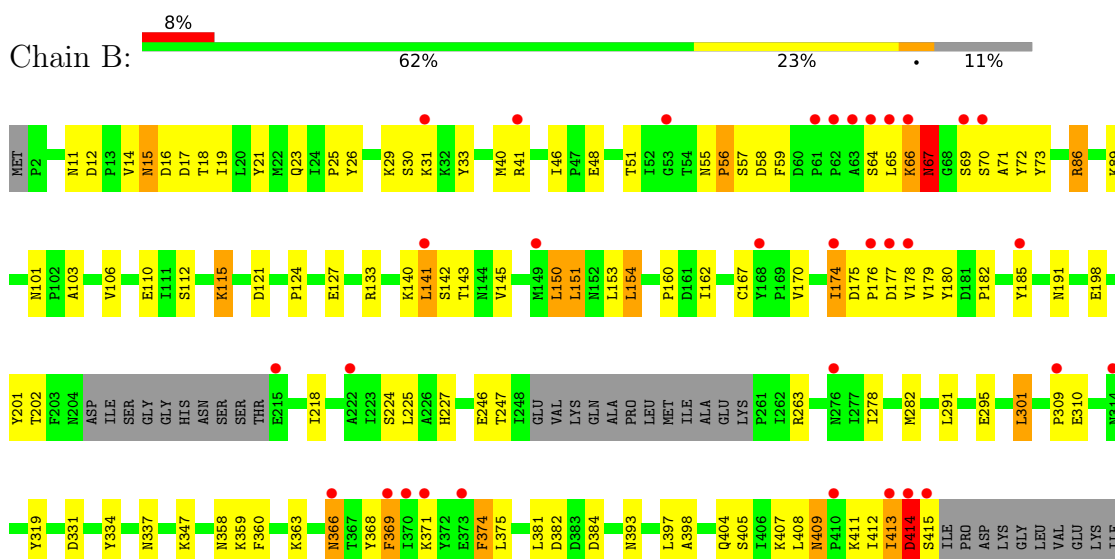
### 3 Residue-property plots i

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: Botulinum neurotoxin type F



- Molecule 1: Botulinum neurotoxin type F



VAL  
LYS  
PHE  
CYS  
LYS  
SER  
VAL  
ILE  
PRO  
ARG  
LYS  
GLY  
THR  
LYS

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	173.40Å 53.24Å 113.87Å 90.00° 119.17° 90.00°	Depositor
Resolution (Å)	26.05 – 1.80 26.05 – 1.80	Depositor EDS
% Data completeness (in resolution range)	76.6 (26.05-1.80) 76.7 (26.05-1.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.28 (at 1.80Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.241 , 0.269 0.240 , 0.267	Depositor DCC
$R_{free}$ test set	1974 reflections (2.83%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.7	Xtrriage
Anisotropy	0.374	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 54.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6586	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.57% 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: ZN, CD

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	0.34	0/3245	0.62	0/4407
1	B	0.36	0/3245	0.61	0/4407
All	All	0.35	0/6490	0.61	0/8814

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	3168	0	3090	155	0
1	B	3168	0	3090	124	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	5	0	0	0	0
3	B	2	0	0	0	0
4	A	109	0	0	14	0
4	B	132	0	0	13	0
All	All	6586	0	6180	274	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 22.

The worst 5 of 274 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:B:86:ARG:HH11	1:B:86:ARG:HB2	1.04	1.11
1:A:174:ILE:HG22	1:A:176:PRO:HD2	1.30	1.09
1:A:365:ARG:HD3	1:A:370:ILE:HD13	1.44	0.97
1:B:363:LYS:H	1:B:404:GLN:HE22	1.04	0.96
1:A:67:ASN:HD21	1:A:73:TYR:H	1.05	0.96

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	386/439 (88%)	373 (97%)	11 (3%)	2 (0%)	29	15
1	B	386/439 (88%)	358 (93%)	20 (5%)	8 (2%)	7	1
All	All	772/878 (88%)	731 (95%)	31 (4%)	10 (1%)	12	3

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	67	ASN
1	B	69	SER
1	B	374	PHE
1	B	413	ILE
1	B	414	ASP

### 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	354/395 (90%)	334 (94%)	20 (6%)	21	8
1	B	354/395 (90%)	334 (94%)	20 (6%)	21	8
All	All	708/790 (90%)	668 (94%)	40 (6%)	21	8

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	150	LEU
1	B	366	ASN
1	B	151	LEU
1	B	185	TYR
1	B	407	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	99	ASN
1	B	329	ASN
1	B	122	HIS
1	B	339	ASN
1	A	314	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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 [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	392/439 (89%)	0.46	27 (6%) 16 13	16, 28, 41, 52	0
1	B	392/439 (89%)	0.49	33 (8%) 11 8	13, 26, 43, 51	0
All	All	784/878 (89%)	0.47	60 (7%) 13 10	13, 27, 42, 52	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	369	PHE	13.0
1	A	414	ASP	12.1
1	A	415	SER	9.4
1	A	2	PRO	8.2
1	B	414	ASP	7.9

### 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 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( $\text{\AA}^2$ )	Q<0.9
3	CD	A	5504	1/1	0.98	0.02	45,45,45,45	0
3	CD	A	5506	1/1	0.98	0.08	39,39,39,39	0
3	CD	B	5505	1/1	0.98	0.11	53,53,53,53	0
2	ZN	B	2437	1/1	0.99	0.14	17,17,17,17	0
3	CD	A	5507	1/1	0.99	0.05	29,29,29,29	0
3	CD	B	5503	1/1	0.99	0.04	46,46,46,46	0
2	ZN	A	1437	1/1	0.99	0.16	18,18,18,18	0
3	CD	A	5502	1/1	1.00	0.02	32,32,32,32	0
3	CD	A	5500	1/1	1.00	0.06	21,21,21,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.