



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 01:08 pm BST

PDB ID : 3N7J  
Title : Crystal structure of botulinum neurotoxin serotype D binding domain  
Authors : Fu, Z.; Baldwin, M.R.; Karalewitz, A.; Kroken, A.; Barbieri, J.T.; Kim, J.-J.P.  
Deposited on : 2010-05-27  
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

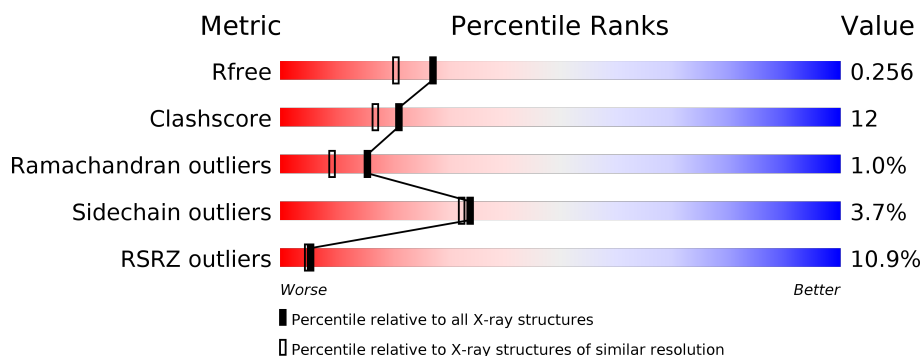
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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	415	<div> <div>11%</div> <div>72%</div> <div>23%</div> <div>..</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3514 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 D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	405	Total	C	N	O	S	0	0	0
			3335	2137	549	640	9			

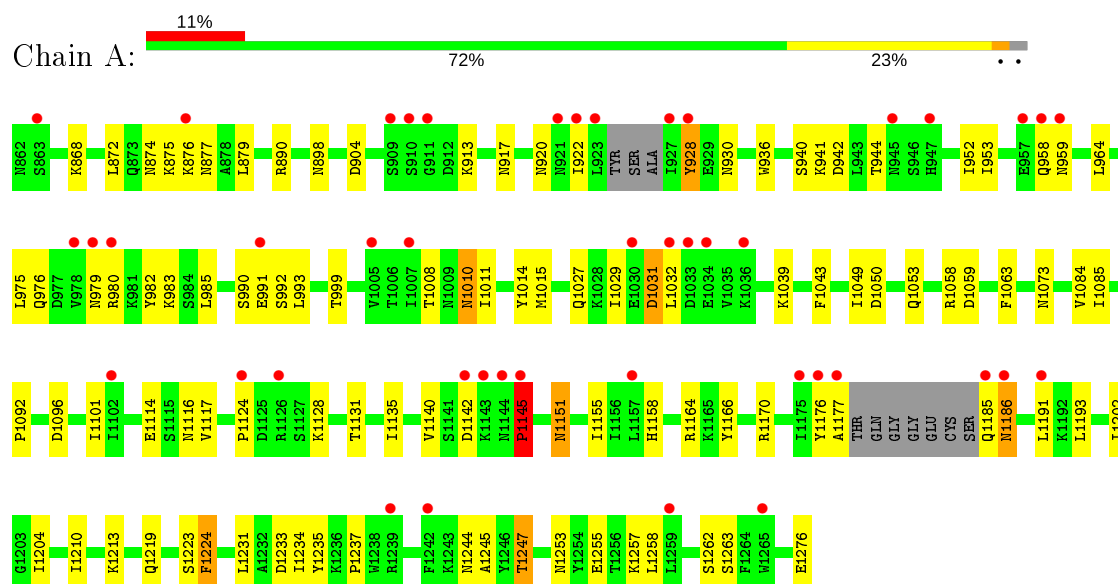
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	179	Total	O	0	0
			179	179		

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

These plots are drawn for all protein, RNA and DNA 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 D



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.04Å 94.15Å 94.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.80 – 2.00 30.01 – 2.00	Depositor EDS
% Data completeness (in resolution range)	89.4 (29.80-2.00) 89.5 (30.01-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.93 (at 2.00Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.222 , 0.256 0.222 , 0.256	Depositor DCC
$R_{free}$ test set	3310 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.5	Xtriage
Anisotropy	0.793	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 48.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.022 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3514	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.03% 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	0.34	0/3404	0.62	0/4609

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	3335	0	3289	80	0
2	A	179	0	0	1	0
All	All	3514	0	3289	80	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.

All (80) 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:A:944:THR:HG21	1:A:999:THR:HG23	1.48	0.94
1:A:958:GLN:HG2	1:A:959:ASN:H	1.33	0.91
1:A:1128:LYS:HD3	1:A:1276:GLU:OE2	1.88	0.74
1:A:920:ASN:ND2	1:A:922:ILE:HG22	2.05	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1073:ASN:HD21	1:A:1213:LYS:NZ	1.88	0.71
1:A:1234:ILE:HD11	1:A:1244:ASN:HB3	1.72	0.70
1:A:1015:MET:SD	1:A:1029:ILE:HD11	2.33	0.69
1:A:1231:LEU:HD11	1:A:1257:LYS:HG2	1.75	0.68
1:A:1050:ASP:H	1:A:1053:GLN:NE2	1.92	0.67
1:A:944:THR:HG21	1:A:999:THR:CG2	2.24	0.67
1:A:983:LYS:HB2	1:A:1032:LEU:HD11	1.81	0.63
1:A:940:SER:O	1:A:944:THR:HG22	1.99	0.63
1:A:936:TRP:HB2	1:A:1058:ARG:HG2	1.80	0.62
1:A:1010:ASN:C	1:A:1010:ASN:HD22	2.05	0.60
1:A:920:ASN:HD22	1:A:922:ILE:HG22	1.67	0.60
1:A:958:GLN:H	1:A:958:GLN:CD	2.06	0.59
1:A:1176:TYR:O	1:A:1177:ALA:HB3	2.02	0.58
1:A:1231:LEU:HD21	1:A:1263:SER:HB3	1.85	0.58
1:A:1010:ASN:HD21	1:A:1014:TYR:H	1.53	0.57
1:A:1231:LEU:CD2	1:A:1263:SER:HB3	2.35	0.56
1:A:1073:ASN:ND2	1:A:1213:LYS:NZ	2.53	0.56
1:A:958:GLN:HG2	1:A:959:ASN:N	2.13	0.55
1:A:1050:ASP:H	1:A:1053:GLN:HE21	1.54	0.55
1:A:952:ILE:HD11	1:A:964:LEU:HD23	1.88	0.55
1:A:1073:ASN:HD21	1:A:1213:LYS:HZ3	1.52	0.55
1:A:941:LYS:HA	1:A:944:THR:CG2	2.38	0.54
1:A:1015:MET:HB3	1:A:1029:ILE:HD11	1.90	0.54
1:A:958:GLN:H	1:A:958:GLN:NE2	2.05	0.54
1:A:1116:ASN:HD21	1:A:1193:LEU:H	1.55	0.54
1:A:985:LEU:HD11	1:A:1015:MET:HG3	1.89	0.54
1:A:952:ILE:HG13	1:A:953:ILE:HG13	1.89	0.52
1:A:1223:SER:O	1:A:1224:PHE:O	2.28	0.52
1:A:1255:GLU:O	1:A:1258:LEU:HD23	2.09	0.52
1:A:1117:VAL:HG11	1:A:1237:PRO:HB2	1.91	0.51
1:A:1116:ASN:ND2	1:A:1193:LEU:H	2.08	0.51
1:A:1155:ILE:C	1:A:1155:ILE:HD12	2.31	0.50
1:A:1164:ARG:HD3	1:A:1166:TYR:CZ	2.46	0.50
1:A:1210:ILE:HD11	1:A:1219:GLN:CD	2.32	0.50
1:A:1073:ASN:HD21	1:A:1213:LYS:HZ1	1.59	0.50
1:A:1096:ASP:OD2	1:A:1145:PRO:O	2.30	0.49
1:A:1185:GLN:HA	1:A:1253:ASN:HB2	1.94	0.49
1:A:1176:TYR:O	1:A:1177:ALA:CB	2.60	0.49
1:A:1010:ASN:ND2	1:A:1014:TYR:H	2.10	0.49
1:A:1231:LEU:HD23	1:A:1263:SER:CA	2.42	0.49
1:A:904:ASP:OD1	1:A:904:ASP:C	2.51	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:941:LYS:HA	1:A:944:THR:HG22	1.94	0.49
1:A:913:LYS:HB2	1:A:1043:PHE:O	2.13	0.49
1:A:1245:ALA:O	1:A:1247:THR:HG22	2.13	0.48
1:A:952:ILE:HD11	1:A:964:LEU:CD2	2.42	0.48
1:A:1124:PRO:HD2	1:A:1131:THR:HG21	1.95	0.48
1:A:1140:VAL:CG2	1:A:1204:ILE:HD12	2.43	0.48
1:A:1049:ILE:HB	1:A:1053:GLN:HE21	1.79	0.48
1:A:1031:ASP:N	1:A:1031:ASP:OD2	2.48	0.47
1:A:1231:LEU:HD22	1:A:1262:SER:OG	2.14	0.47
1:A:975:LEU:HD23	1:A:1029:ILE:HD13	1.97	0.46
1:A:1085:ILE:HD11	1:A:1151:ASN:HB3	1.96	0.46
1:A:1170:ARG:O	1:A:1202:ILE:HD11	2.15	0.46
1:A:930:ASN:OD1	1:A:1011:ILE:HG23	2.14	0.46
1:A:1073:ASN:ND2	1:A:1213:LYS:HZ3	2.13	0.46
1:A:1151:ASN:ND2	1:A:1151:ASN:O	2.48	0.45
1:A:976:GLN:HB3	1:A:982:TYR:HB3	1.98	0.45
1:A:1140:VAL:HG21	1:A:1204:ILE:HD12	1.99	0.45
1:A:1151:ASN:HD22	1:A:1151:ASN:N	2.15	0.45
1:A:1158:HIS:HE1	2:A:165:HOH:O	2.00	0.44
1:A:1015:MET:HG2	1:A:1027:GLN:HB3	2.00	0.44
1:A:1101:ILE:HD12	1:A:1135:ILE:HD12	1.98	0.44
1:A:1058:ARG:HG3	1:A:1059:ASP:OD2	2.18	0.44
1:A:1233:ASP:HB3	1:A:1235:TYR:CE1	2.52	0.44
1:A:979:ASN:O	1:A:980:ARG:HB2	2.18	0.44
1:A:875:LYS:HA	1:A:875:LYS:HD2	1.65	0.43
1:A:992:SER:O	1:A:993:LEU:HB2	2.19	0.43
1:A:876:LYS:HG2	1:A:877:ASN:N	2.32	0.43
1:A:1008:THR:O	1:A:1015:MET:HA	2.19	0.42
1:A:1073:ASN:ND2	1:A:1213:LYS:HZ1	2.16	0.42
1:A:1084:VAL:HG11	1:A:1092:PRO:HB3	2.02	0.42
1:A:928:TYR:O	1:A:1011:ILE:HG22	2.19	0.41
1:A:990:SER:O	1:A:991:GLU:C	2.58	0.41
1:A:1039:LYS:HA	1:A:1039:LYS:HD3	1.89	0.41
1:A:890:ARG:HH11	1:A:917:ASN:HD21	1.67	0.40
1:A:868:LYS:HG2	1:A:1063:PHE:CE1	2.57	0.40

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	399/415 (96%)	364 (91%)	31 (8%)	4 (1%)	15 9

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1145	PRO
1	A	1224	PHE
1	A	1186	ASN
1	A	1142	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	380/387 (98%)	366 (96%)	14 (4%)	34 32

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	872	LEU
1	A	874	ASN
1	A	879	LEU
1	A	898	ASN
1	A	928	TYR
1	A	942	ASP
1	A	1010	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1031	ASP
1	A	1114	GLU
1	A	1145	PRO
1	A	1151	ASN
1	A	1186	ASN
1	A	1191	LEU
1	A	1247	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	898	ASN
1	A	920	ASN
1	A	948	ASN
1	A	958	GLN
1	A	968	ASN
1	A	970	ASN
1	A	1000	ASN
1	A	1010	ASN
1	A	1025	GLN
1	A	1053	GLN
1	A	1061	ASN
1	A	1073	ASN
1	A	1116	ASN
1	A	1151	ASN
1	A	1185	GLN
1	A	1186	ASN
1	A	1199	ASN
1	A	1219	GLN

### 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 [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	405/415 (97%)	0.67	44 (10%) 5 5	19, 40, 72, 90	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	928	TYR	7.8
1	A	1177	ALA	7.3
1	A	1176	TYR	5.9
1	A	980	ARG	5.8
1	A	1185	GLN	5.1
1	A	979	ASN	5.0
1	A	1034	GLU	4.9
1	A	927	ILE	4.7
1	A	957	GLU	4.4
1	A	1242	PHE	4.4
1	A	978	VAL	3.7
1	A	910	SER	3.7
1	A	1145	PRO	3.6
1	A	1143	LYS	3.5
1	A	876	LYS	3.4
1	A	1126	ARG	3.3
1	A	922	ILE	3.3
1	A	1030	GLU	3.1
1	A	923	LEU	3.1
1	A	1036	LYS	3.0
1	A	1144	ASN	3.0
1	A	959	ASN	3.0
1	A	909	SER	3.0
1	A	958	GLN	2.9
1	A	1032	LEU	2.8
1	A	1102	ILE	2.8
1	A	921	ASN	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1239	ARG	2.7
1	A	1142	ASP	2.7
1	A	1175	ILE	2.7
1	A	911	GLY	2.7
1	A	1033	ASP	2.5
1	A	1007	ILE	2.3
1	A	1157	LEU	2.3
1	A	945	ASN	2.3
1	A	1191	LEU	2.2
1	A	1186	ASN	2.2
1	A	1124	PRO	2.2
1	A	1005	VAL	2.2
1	A	947	HIS	2.1
1	A	1259	LEU	2.1
1	A	1265	TRP	2.1
1	A	991	GLU	2.1
1	A	863	SER	2.1

## 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 carbohydrates 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.