WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2024 – 07:35 PM EST

PDB ID	:	3R4U
Title	:	Cell entry of botulinum neurotoxin type C is dependent upon interaction with two ganglioside molecules
Authors	:	Strotmeier, J.; Gu, S.; Jutzi, S.; Mahrhold, S.; Zhou, J.; Pich, A.; Bigalke, H.; Rummel, A.; Jin, R.; Binz, T.
Deposited on	:	2011-03-17
Resolution	:	2.20 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

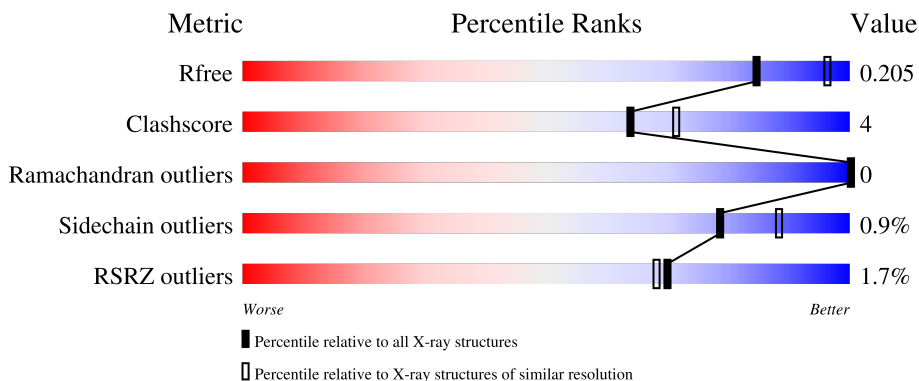|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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	Percentile Ranks	Value
Rfree		0.205
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		0.9%
RSRZ outliers		1.7%

Worse .... *Better*

▌ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	443	% 87% 8% 5%
1	B	443	2% 86% 8% 5%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 7641 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 C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	0	0	0
			3465	2206	580	666	13			
1	B	422	Total	C	N	O	S	0	0	0
			3465	2206	580	666	13			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	861	MET	-	expression tag	UNP P18640
A	862	ARG	-	expression tag	UNP P18640
A	863	GLY	-	expression tag	UNP P18640
A	864	SER	-	expression tag	UNP P18640
A	865	MET	-	expression tag	UNP P18640
A	866	ALA	-	expression tag	UNP P18640
A	1292	PRO	-	expression tag	UNP P18640
A	1293	GLY	-	expression tag	UNP P18640
A	1294	SER	-	expression tag	UNP P18640
A	1295	ALA	-	expression tag	UNP P18640
A	1296	TRP	-	expression tag	UNP P18640
A	1297	SER	-	expression tag	UNP P18640
A	1298	HIS	-	expression tag	UNP P18640
A	1299	PRO	-	expression tag	UNP P18640
A	1300	GLN	-	expression tag	UNP P18640
A	1301	PHE	-	expression tag	UNP P18640
A	1302	GLU	-	expression tag	UNP P18640
A	1303	LYS	-	expression tag	UNP P18640
B	861	MET	-	expression tag	UNP P18640
B	862	ARG	-	expression tag	UNP P18640
B	863	GLY	-	expression tag	UNP P18640
B	864	SER	-	expression tag	UNP P18640
B	865	MET	-	expression tag	UNP P18640
B	866	ALA	-	expression tag	UNP P18640
B	1292	PRO	-	expression tag	UNP P18640

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1293	GLY	-	expression tag	UNP P18640
B	1294	SER	-	expression tag	UNP P18640
B	1295	ALA	-	expression tag	UNP P18640
B	1296	TRP	-	expression tag	UNP P18640
B	1297	SER	-	expression tag	UNP P18640
B	1298	HIS	-	expression tag	UNP P18640
B	1299	PRO	-	expression tag	UNP P18640
B	1300	GLN	-	expression tag	UNP P18640
B	1301	PHE	-	expression tag	UNP P18640
B	1302	GLU	-	expression tag	UNP P18640
B	1303	LYS	-	expression tag	UNP P18640

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	352	Total 352	O 352	0	0
2	B	359	Total 359	O 359	0	0

3 Residue-property plots ⓘ

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 C1

Chain A:

• Molecule 1: Botulinum neurotoxin type C1

Chain B:
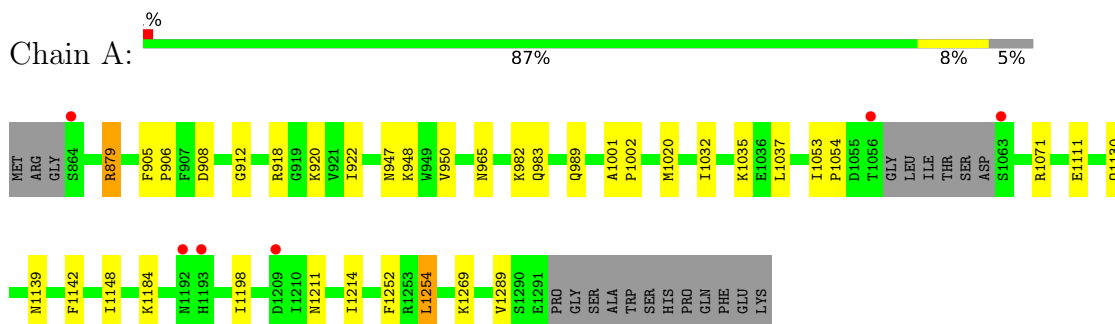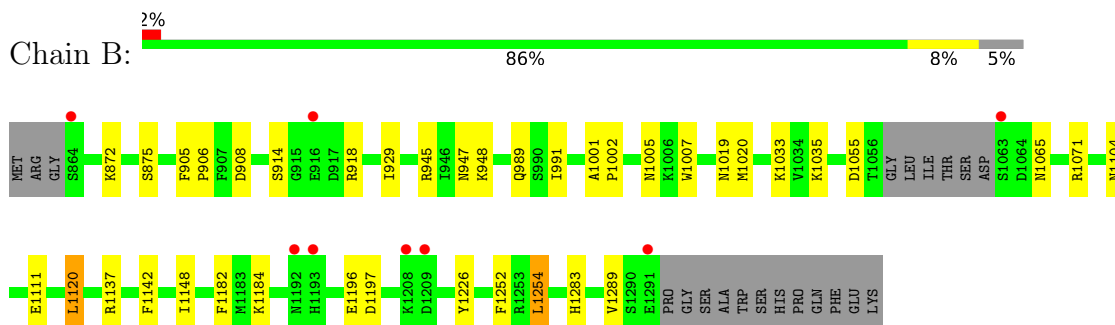
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	98.02Å 76.50Å 106.94Å 90.00° 116.27° 90.00°	Depositor
Resolution (Å)	44.10 – 2.20 44.10 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.0 (44.10-2.20) 98.9 (44.10-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.21 (at 2.20Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.171 , 0.210 0.166 , 0.205	Depositor DCC
R_{free} test set	3617 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	30.9	Xtriage
Anisotropy	0.482	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 41.1	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.009 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7641	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.36% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

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	0.39	0/3542	0.55	0/4794
1	B	0.41	0/3542	0.55	0/4794
All	All	0.40	0/7084	0.55	0/9588

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	3465	0	3337	31	0
1	B	3465	0	3337	26	0
2	A	352	0	0	1	0
2	B	359	0	0	1	0
All	All	7641	0	6674	57	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 4.

All (57) 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:989:GLN:HE22	1:A:1035:LYS:H	1.14	0.95

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:989:GLN:HE22	1:B:1035:LYS:H	1.18	0.91
1:A:879:ARG:HH11	1:A:879:ARG:HG3	1.48	0.79
1:A:879:ARG:HH11	1:A:879:ARG:CG	1.99	0.75
1:A:920:LYS:HE2	1:A:922:ILE:HD11	1.73	0.71
1:A:1020:MET:HE3	1:A:1032:ILE:HD11	1.72	0.69
1:A:948:LYS:HE3	1:A:1002:PRO:O	1.93	0.69
1:B:989:GLN:HG2	1:B:1020:MET:HE1	1.76	0.68
1:B:1226:TYR:OH	1:B:1283:HIS:HD2	1.78	0.66
1:B:1111:GLU:O	1:B:1289:VAL:HG13	1.96	0.66
1:A:1252:PHE:HB3	1:A:1254:LEU:HD13	1.76	0.65
1:B:1289:VAL:CG2	1:B:1289:VAL:O	2.45	0.63
1:B:1007:TRP:HE1	1:B:1104:ASN:HD21	1.50	0.59
1:B:1148:ILE:C	1:B:1148:ILE:HD12	2.24	0.57
1:B:1289:VAL:O	1:B:1289:VAL:HG22	2.07	0.55
1:A:879:ARG:HG3	1:A:879:ARG:NH1	2.19	0.55
1:B:872:LYS:HD3	1:B:875:SER:HB2	1.90	0.54
1:A:989:GLN:HE22	1:A:1035:LYS:N	1.96	0.52
1:A:965:ASN:HD22	1:A:982:LYS:HE2	1.76	0.51
1:A:879:ARG:NH1	2:A:284:HOH:O	2.44	0.50
1:A:1130:GLN:HE21	1:A:1184:LYS:NZ	2.09	0.49
1:A:1198:ILE:HD13	1:A:1269:LYS:HD3	1.93	0.49
1:B:1020:MET:HE3	1:B:1020:MET:HB3	1.68	0.48
1:B:1252:PHE:HB3	1:B:1254:LEU:HD13	1.95	0.48
1:A:1001:ALA:N	1:A:1002:PRO:CD	2.77	0.48
1:B:989:GLN:HE22	1:B:1035:LYS:N	2.00	0.48
1:A:905:PHE:CG	1:A:906:PRO:HA	2.49	0.47
1:A:908:ASP:HA	1:A:1071:ARG:HB3	1.95	0.47
1:A:989:GLN:NE2	1:A:1035:LYS:HG2	2.29	0.47
1:B:1001:ALA:N	1:B:1002:PRO:CD	2.77	0.47
1:B:914:SER:OG	1:B:1065:ASN:HB2	2.16	0.46
1:A:989:GLN:OE1	1:A:1032:ILE:HD13	2.14	0.46
1:A:1211:ASN:O	1:A:1214:ILE:HG12	2.15	0.46
1:B:947:ASN:HA	1:B:948:LYS:HA	1.73	0.46
1:B:1019:ASN:HD22	1:B:1033:LYS:HA	1.81	0.45
1:B:945:ARG:HD2	1:B:1005:ASN:O	2.16	0.45
1:A:1148:ILE:HD12	1:A:1148:ILE:C	2.36	0.45
1:B:905:PHE:CG	1:B:906:PRO:HA	2.51	0.45
1:A:948:LYS:HG2	1:A:950:VAL:HG13	1.98	0.45
1:B:918:ARG:HG2	2:B:519:HOH:O	2.17	0.45
1:A:1198:ILE:CD1	1:A:1269:LYS:HD3	2.47	0.45
1:A:947:ASN:HA	1:A:948:LYS:HA	1.72	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1111:GLU:O	1:A:1289:VAL:HG13	2.17	0.44
1:B:991:ILE:HD13	1:B:1020:MET:HG2	2.00	0.43
1:B:1120:LEU:HD12	1:B:1120:LEU:HA	1.81	0.43
1:A:983:GLN:HB2	1:A:1037:LEU:HD22	1.99	0.43
1:B:1002:PRO:HG2	1:B:1142:PHE:CE2	2.53	0.43
1:A:1111:GLU:HB3	1:A:1289:VAL:HG11	2.00	0.43
1:B:1182:PHE:CZ	1:B:1184:LYS:HG2	2.54	0.42
1:A:1001:ALA:HB3	1:A:1002:PRO:HD3	2.01	0.42
1:B:908:ASP:HA	1:B:1071:ARG:HB3	2.02	0.42
1:A:1002:PRO:HG2	1:A:1142:PHE:CE2	2.54	0.41
1:A:1053:ILE:HA	1:A:1054:PRO:HD3	1.74	0.41
1:A:912:GLY:O	1:A:918:ARG:HG3	2.20	0.41
1:A:1111:GLU:CB	1:A:1289:VAL:HG11	2.50	0.41
1:B:929:ILE:HD12	1:B:929:ILE:HA	1.93	0.41
1:B:1196:GLU:O	1:B:1197:ASP:HB2	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	418/443 (94%)	401 (96%)	17 (4%)	0	100	100
1	B	418/443 (94%)	405 (97%)	13 (3%)	0	100	100
All	All	836/886 (94%)	806 (96%)	30 (4%)	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	386/403 (96%)	383 (99%)	3 (1%)	81	90
1	B	386/403 (96%)	382 (99%)	4 (1%)	76	86
All	All	772/806 (96%)	765 (99%)	7 (1%)	78	88

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	879	ARG
1	A	1139	ASN
1	A	1254	LEU
1	B	1055	ASP
1	B	1120	LEU
1	B	1137	ARG
1	B	1254	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (22) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	932	ASN
1	A	965	ASN
1	A	989	GLN
1	A	1086	ASN
1	A	1130	GLN
1	A	1138	ASN
1	A	1222	ASN
1	B	877	GLN
1	B	890	ASN
1	B	932	ASN
1	B	965	ASN
1	B	989	GLN
1	B	999	ASN
1	B	1019	ASN
1	B	1093	GLN
1	B	1104	ASN
1	B	1175	ASN
1	B	1193	HIS
1	B	1222	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1262	ASN
1	B	1270	GLN
1	B	1283	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 95th 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(Å²)	Q<0.9
1	A	422/443 (95%)	-0.57	6 (1%)	75	73	20, 32, 58, 91	0
1	B	422/443 (95%)	-0.57	8 (1%)	66	65	20, 32, 58, 92	0
All	All	844/886 (95%)	-0.57	14 (1%)	70	68	20, 32, 58, 92	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1192	ASN	4.0
1	A	1056	THR	3.6
1	B	864	SER	3.0
1	B	1193	HIS	2.8
1	A	864	SER	2.5
1	A	1209	ASP	2.5
1	A	1192	ASN	2.5
1	A	1193	HIS	2.4
1	B	1208	LYS	2.4
1	B	1291	GLU	2.2
1	B	1063	SER	2.2
1	B	1209	ASP	2.2
1	B	916	GLU	2.1
1	A	1063	SER	2.1

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 monosaccharides 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.