



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 25, 2020 – 06:45 PM EST

PDB ID : 6UI1
Title : Crystal structure of BoNT/A-LCHn domain in complex with VHH ciA-D12, ciA-B5, and ciA-H7
Authors : Lam, K.; Jin, R.
Deposited on : 2019-09-29
Resolution : 2.20 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

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 : 2.8
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
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) : 2.8

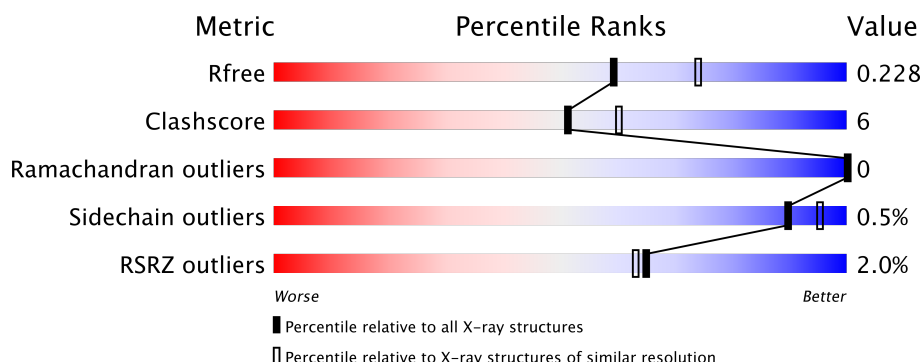
1 Overall quality at a glance

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	4343 (2.20-2.20)
Clashscore	122126	5027 (2.20-2.20)
Ramachandran outliers	120053	4952 (2.20-2.20)
Sidechain outliers	120020	4953 (2.20-2.20)
RSRZ outliers	108989	4245 (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 on 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 $\leq 5\%$. 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	873	<div> <div>2%</div> <div>85%12%</div> </div>
2	B	123	<div> <div>2%</div> <div>81%11%7%</div> </div>
3	C	134	<div> <div>%</div> <div>84%7%8%</div> </div>
4	D	130	<div> <div></div> <div>75%16%9%</div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10088 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 BoNT/A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	851	Total	C	N	O	S	0	0	0
			6868	4420	1111	1317	20			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q7B8V4
A	0	PRO	-	expression tag	UNP Q7B8V4
A	224	GLN	GLU	conflict	UNP Q7B8V4
A	363	ALA	ARG	conflict	UNP Q7B8V4
A	366	PHE	TYR	conflict	UNP Q7B8V4

- Molecule 2 is a protein called ciA-H7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	114	Total	C	N	O	S	0	0	0
			855	529	152	169	5			

- Molecule 3 is a protein called ciA-D12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	123	Total	C	N	O	S	0	0	0
			933	590	159	179	5			

- Molecule 4 is a protein called ciA-B5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	118	Total	C	N	O	S	0	0	0
			908	568	161	174	5			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	371	Total 371	O 371	0	0
5	B	59	Total 59	O 59	0	0
5	C	49	Total 49	O 49	0	0
5	D	45	Total 45	O 45	0	0

- Molecule 1: BoNT/A



GLY	PRO	LEU	GLY	SER	GLN	VAL	Q3	E6	C22	A23	P24	L28	W42	Y43	L59	A64	D65	S66	V67	R70	T81	P91	E92	D93	E102	H105	G110	V115	K120	THR	PRO	LYS	PRO	GLN
-----	-----	-----	-----	-----	-----	-----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	141.04Å 89.36Å 143.69Å 90.00° 118.93° 90.00°	Depositor
Resolution (Å)	37.16 – 2.20 37.16 – 2.20	Depositor EDS
% Data completeness (in resolution range)	94.8 (37.16-2.20) 94.8 (37.16-2.20)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.42 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.9_1692, PHENIX 1.9_1692	Depositor
R, R_{free}	0.181 , 0.229 0.182 , 0.228	Depositor DCC
R_{free} test set	3774 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	29.8	Xtriage
Anisotropy	0.225	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 36.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.016 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	10088	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.23% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.45	0/7013	0.54	0/9500
2	B	0.41	0/869	0.58	0/1170
3	C	0.41	0/954	0.55	0/1293
4	D	0.43	0/931	0.55	0/1265
All	All	0.44	0/9767	0.55	0/13228

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	6868	0	6778	88	0
2	B	855	0	835	12	0
3	C	933	0	905	11	0
4	D	908	0	872	13	0
5	A	371	0	0	7	0
5	B	59	0	0	2	0
5	C	49	0	0	2	0
5	D	45	0	0	2	0
All	All	10088	0	9390	116	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 6.

The worst 5 of 116 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:482:THR:HG22	1:A:484:ASP:H	1.10	1.08
1:A:67:GLN:HE22	1:A:537:PHE:H	1.19	0.84
1:A:124:ASP:HB3	2:B:44:ARG:HA	1.60	0.83
1:A:482:THR:HG22	1:A:484:ASP:N	1.92	0.82
1:A:67:GLN:HE22	1:A:537:PHE:N	1.80	0.78

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	845/873 (97%)	823 (97%)	22 (3%)	0	100	100
2	B	112/123 (91%)	111 (99%)	1 (1%)	0	100	100
3	C	121/134 (90%)	121 (100%)	0	0	100	100
4	D	116/130 (89%)	113 (97%)	3 (3%)	0	100	100
All	All	1194/1260 (95%)	1168 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	763/786 (97%)	760 (100%)	3 (0%)	92	96
2	B	91/98 (93%)	90 (99%)	1 (1%)	76	87
3	C	97/107 (91%)	97 (100%)	0	100	100
4	D	99/110 (90%)	98 (99%)	1 (1%)	78	88
All	All	1050/1101 (95%)	1045 (100%)	5 (0%)	90	95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	GLN
1	A	165	CYS
1	A	550	MET
2	B	28	ASP
4	D	59	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	67	GLN
1	A	722	ASN
1	A	774	ASN
3	C	30	ASN

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 carbohydrates 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, 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(Å ²)	Q<0.9
1	A	851/873 (97%)	-0.25	19 (2%) 62 59	17, 32, 57, 93	0
2	B	114/123 (92%)	-0.27	3 (2%) 56 53	23, 33, 50, 72	0
3	C	123/134 (91%)	-0.45	2 (1%) 72 70	24, 35, 51, 79	0
4	D	118/130 (90%)	-0.43	0 100 100	26, 34, 52, 62	0
All	All	1206/1260 (95%)	-0.29	24 (1%) 65 63	17, 33, 54, 93	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	864	SER	6.9
2	B	10	GLY	4.2
1	A	647	LEU	4.0
1	A	865	THR	3.4
2	B	1	GLN	3.4

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.