Full wwPDB X-ray Structure Validation Report (i)

Dec 10, 2023 – 04:50 am GMT

PDB ID	:	2VUA
Title	:	Crystal Structure of the Botulinum Neurotoxin Serotype A binding domain
Authors	:	Stenmark, P.; Dupuy, J.; Stevens, R.C.
Deposited on	:	2008-05-22
Resolution	:	1.70 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

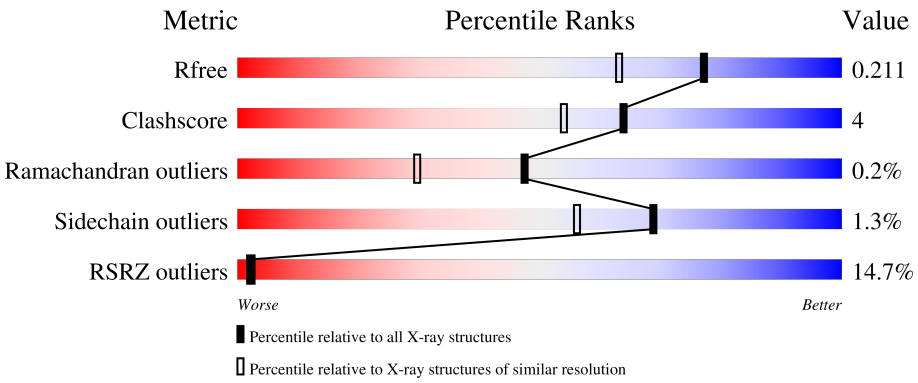
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

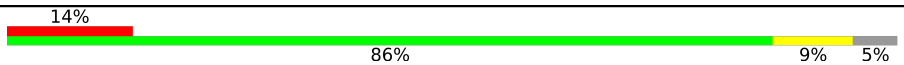
The reported resolution of this entry is 1.70 Å.

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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	444	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3825 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 A HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	423	3514	2238	607	655	14	0	4	0

- Molecule 2 is water.

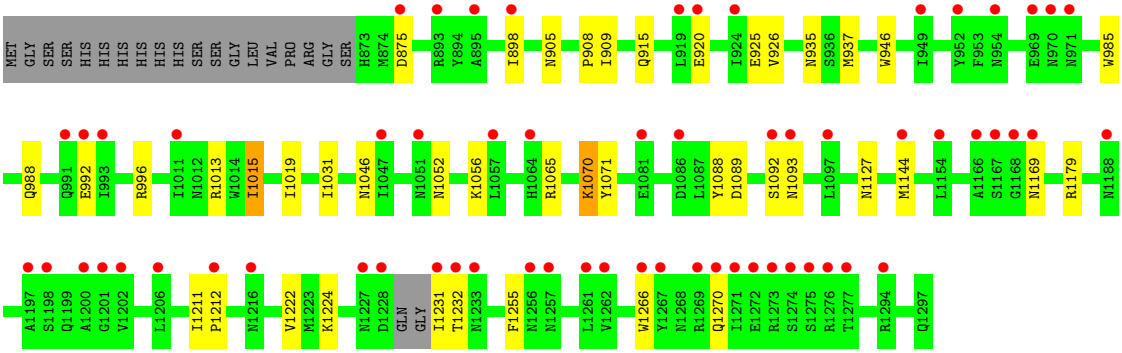
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	311	311	311	0	0

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 A HEAVY CHAIN

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	73.00Å 114.50Å 105.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.73 – 1.70 19.73 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.0 (19.73-1.70) 99.0 (19.73-1.70)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.55 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.168 , 0.205 0.178 , 0.211	Depositor DCC
R_{free} test set	2469 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	23.7	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 56.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3825	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.70% 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 (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	0.66	1/3599 (0.0%)	0.77	1/4861 (0.0%)

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
1	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1070	LYS	CE-NZ	5.18	1.62	1.49

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	996	ARG	NE-CZ-NH1	5.65	123.13	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1255	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	3514	0	3479	29	0
2	A	311	0	0	8	0
All	All	3825	0	3479	29	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 (29) 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:1013[A]:ARG:HH12	1:A:1231:ILE:HD11	0.97	1.07
1:A:1013[A]:ARG:NH1	1:A:1231:ILE:HD11	1.80	0.97
1:A:875:ASP:OD2	2:A:2003:HOH:O	1.91	0.87
1:A:1015:ILE:HG22	2:A:2089:HOH:O	1.83	0.78
1:A:1013[A]:ARG:HH12	1:A:1231:ILE:CD1	1.90	0.76
1:A:1232:THR:HG21	2:A:2094:HOH:O	1.94	0.68
1:A:925[B]:GLU:CD	1:A:1052:ASN:HD22	1.98	0.66
1:A:1089:ASP:O	1:A:1093:ASN:N	2.30	0.64
1:A:1092:SER:O	2:A:2143:HOH:O	2.16	0.60
1:A:1070:LYS:HE2	1:A:1071:TYR:CE2	2.37	0.59
1:A:988:GLN:HE21	1:A:992:GLU:HA	1.69	0.58
1:A:1070:LYS:HD3	2:A:2125:HOH:O	2.03	0.58
1:A:898:ILE:HD12	1:A:926:VAL:HG22	1.87	0.56
1:A:1179:ARG:HG2	1:A:1222:VAL:HG22	1.88	0.55
1:A:1070:LYS:HE2	1:A:1071:TYR:CZ	2.42	0.53
1:A:1015:ILE:HD13	1:A:1031:ILE:HG23	1.92	0.51
1:A:920:GLU:HA	1:A:1056:LYS:NZ	2.27	0.50
1:A:905:ASN:HB3	1:A:915:GLN:HB3	1.97	0.47
1:A:1088:TYR:O	1:A:1092:SER:OG	2.15	0.46
1:A:1232:THR:CG2	2:A:2094:HOH:O	2.58	0.46
1:A:985:TRP:CD2	1:A:1019:ILE:HG21	2.51	0.45
1:A:937:MET:HG3	1:A:1046:ASN:HD22	1.82	0.45
1:A:1070:LYS:CD	2:A:2125:HOH:O	2.65	0.42
1:A:946:TRP:HB2	1:A:1070:LYS:HG2	2.02	0.42
1:A:1266:TRP:O	1:A:1270:GLN:HG2	2.19	0.42
1:A:908:PRO:HG2	1:A:909:ILE:HD12	2.02	0.41
1:A:1211:ILE:N	1:A:1212:PRO:CD	2.84	0.41
1:A:1224:LYS:NZ	2:A:2258:HOH:O	2.54	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	423/444 (95%)	405 (96%)	17 (4%)	1 (0%)	47 30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1127	ASN

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	394/408 (97%)	389 (99%)	5 (1%)	69 56

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

Mol	Chain	Res	Type
1	A	935	ASN
1	A	1015	ILE
1	A	1065	ARG
1	A	1144	MET
1	A	1169	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	930	ASN
1	A	935	ASN
1	A	988	GLN
1	A	1046	ASN
1	A	1216	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)

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	423/444 (95%)	1.03	62 (14%) 2 2	13, 22, 37, 46	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1271	ILE	10.1
1	A	1256	ASN	8.7
1	A	1231	ILE	8.2
1	A	1272	GLU	7.5
1	A	1166	ALA	5.5
1	A	1093	ASN	5.4
1	A	1200	ALA	5.4
1	A	1274	SER	5.4
1	A	1275	SER	4.7
1	A	1168	GLY	4.6
1	A	875	ASP	4.4
1	A	1154	LEU	4.4
1	A	1276	ARG	4.3
1	A	1294	ARG	4.3
1	A	1273[A]	ARG	4.1
1	A	1227	ASN	4.0
1	A	1198	SER	4.0
1	A	1232	THR	3.9
1	A	1266	TRP	3.9
1	A	1144	MET	3.9
1	A	1047	ILE	3.8
1	A	970	ASN	3.8
1	A	993	ILE	3.7
1	A	1216	ASN	3.7
1	A	920	GLU	3.4
1	A	969	GLU	3.2
1	A	1051	ASN	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	954	ASN	2.9
1	A	1169	ASN	2.9
1	A	1228	ASP	2.9
1	A	1201	GLY	2.8
1	A	1257	ASN	2.8
1	A	1081	GLU	2.8
1	A	1092	SER	2.7
1	A	1011	ILE	2.6
1	A	1188	ASN	2.6
1	A	1064	HIS	2.6
1	A	1167	SER	2.5
1	A	1277	THR	2.5
1	A	1097	LEU	2.5
1	A	895	ALA	2.4
1	A	991	GLN	2.4
1	A	924	ILE	2.4
1	A	893	ARG	2.3
1	A	1206	LEU	2.3
1	A	1270	GLN	2.3
1	A	919	LEU	2.3
1	A	1261	LEU	2.3
1	A	1197	ALA	2.3
1	A	1269	ARG	2.3
1	A	1267	TYR	2.3
1	A	971	ASN	2.3
1	A	1262	VAL	2.2
1	A	1057	LEU	2.2
1	A	952	TYR	2.2
1	A	1202	VAL	2.2
1	A	1212	PRO	2.2
1	A	992	GLU	2.2
1	A	1233	ASN	2.1
1	A	1086	ASP	2.1
1	A	898	ILE	2.1
1	A	949	ILE	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.