



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 02:54 AM GMT

PDB ID : 4JRA  
Title : CRYSTAL STRUCTURE OF THE BOTULINUM NEUROTOXIN A  
RECEPTOR-BINDING DOMAIN IN COMPLEX WITH THE LUMINAL  
DOMAIN Of SV2C  
Authors : Benoit, R.M.; Frey, D.; Wieser, M.M.; Jaussi, R.; Schertler, G.F.X.; Capitani,  
G.; Kammerer, R.A.  
Deposited on : 2013-03-21  
Resolution : 2.30 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

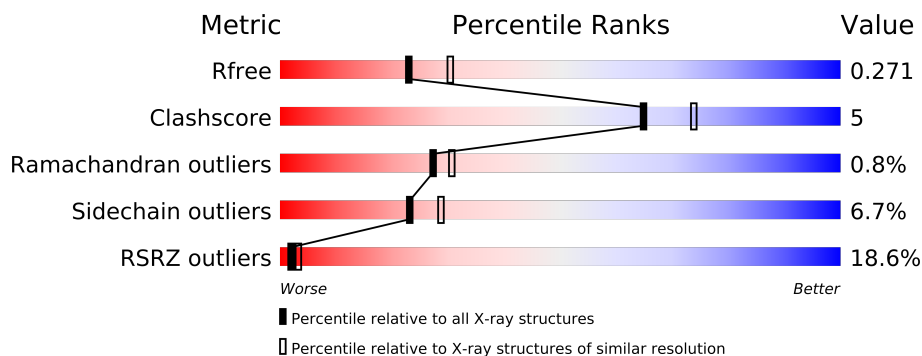
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.30 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))
$R_{free}$	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (2.30-2.30)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	443	
1	B	443	
2	C	136	
2	D	136	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	CL	A	1301	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 16539 atoms, of which 8131 are hydrogens and 0 are deuterium.

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 A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	416	Total	C	H	N	O	S	0	3	0
			6854	2195	3412	593	641	13			
1	B	413	Total	C	H	N	O	S	0	0	0
			6778	2176	3367	585	637	13			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	854	MET	-	EXPRESSION TAG	UNP P10845
A	855	LYS	-	EXPRESSION TAG	UNP P10845
A	856	LYS	-	EXPRESSION TAG	UNP P10845
A	857	HIS	-	EXPRESSION TAG	UNP P10845
A	858	HIS	-	EXPRESSION TAG	UNP P10845
A	859	HIS	-	EXPRESSION TAG	UNP P10845
A	860	HIS	-	EXPRESSION TAG	UNP P10845
A	861	HIS	-	EXPRESSION TAG	UNP P10845
A	862	HIS	-	EXPRESSION TAG	UNP P10845
A	863	GLY	-	EXPRESSION TAG	UNP P10845
A	864	SER	-	EXPRESSION TAG	UNP P10845
A	865	LEU	-	EXPRESSION TAG	UNP P10845
A	866	VAL	-	EXPRESSION TAG	UNP P10845
A	867	PRO	-	EXPRESSION TAG	UNP P10845
A	868	ARG	-	EXPRESSION TAG	UNP P10845
A	869	GLY	-	EXPRESSION TAG	UNP P10845
A	870	SER	-	EXPRESSION TAG	UNP P10845
B	854	MET	-	EXPRESSION TAG	UNP P10845
B	855	LYS	-	EXPRESSION TAG	UNP P10845
B	856	LYS	-	EXPRESSION TAG	UNP P10845
B	857	HIS	-	EXPRESSION TAG	UNP P10845
B	858	HIS	-	EXPRESSION TAG	UNP P10845
B	859	HIS	-	EXPRESSION TAG	UNP P10845
B	860	HIS	-	EXPRESSION TAG	UNP P10845
B	861	HIS	-	EXPRESSION TAG	UNP P10845

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	862	HIS	-	EXPRESSION TAG	UNP P10845
B	863	GLY	-	EXPRESSION TAG	UNP P10845
B	864	SER	-	EXPRESSION TAG	UNP P10845
B	865	LEU	-	EXPRESSION TAG	UNP P10845
B	866	VAL	-	EXPRESSION TAG	UNP P10845
B	867	PRO	-	EXPRESSION TAG	UNP P10845
B	868	ARG	-	EXPRESSION TAG	UNP P10845
B	869	GLY	-	EXPRESSION TAG	UNP P10845
B	870	SER	-	EXPRESSION TAG	UNP P10845

- Molecule 2 is a protein called Synaptic vesicle glycoprotein 2C.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	83	Total	C	H	N	O	S	0	0	0
			1320	444	634	103	134	5			
2	D	93	Total	C	H	N	O	S	0	0	0
			1497	502	718	121	151	5			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	439	MET	-	EXPRESSION TAG	UNP Q496J9
C	440	LYS	-	EXPRESSION TAG	UNP Q496J9
C	441	LYS	-	EXPRESSION TAG	UNP Q496J9
C	442	HIS	-	EXPRESSION TAG	UNP Q496J9
C	443	HIS	-	EXPRESSION TAG	UNP Q496J9
C	444	HIS	-	EXPRESSION TAG	UNP Q496J9
C	445	HIS	-	EXPRESSION TAG	UNP Q496J9
C	446	HIS	-	EXPRESSION TAG	UNP Q496J9
C	447	HIS	-	EXPRESSION TAG	UNP Q496J9
C	448	GLY	-	EXPRESSION TAG	UNP Q496J9
C	449	SER	-	EXPRESSION TAG	UNP Q496J9
C	450	LEU	-	EXPRESSION TAG	UNP Q496J9
C	451	VAL	-	EXPRESSION TAG	UNP Q496J9
C	452	PRO	-	EXPRESSION TAG	UNP Q496J9
C	453	ARG	-	EXPRESSION TAG	UNP Q496J9
C	454	GLY	-	EXPRESSION TAG	UNP Q496J9
C	455	SER	-	EXPRESSION TAG	UNP Q496J9
D	439	MET	-	EXPRESSION TAG	UNP Q496J9
D	440	LYS	-	EXPRESSION TAG	UNP Q496J9
D	441	LYS	-	EXPRESSION TAG	UNP Q496J9
D	442	HIS	-	EXPRESSION TAG	UNP Q496J9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	443	HIS	-	EXPRESSION TAG	UNP Q496J9
D	444	HIS	-	EXPRESSION TAG	UNP Q496J9
D	445	HIS	-	EXPRESSION TAG	UNP Q496J9
D	446	HIS	-	EXPRESSION TAG	UNP Q496J9
D	447	HIS	-	EXPRESSION TAG	UNP Q496J9
D	448	GLY	-	EXPRESSION TAG	UNP Q496J9
D	449	SER	-	EXPRESSION TAG	UNP Q496J9
D	450	LEU	-	EXPRESSION TAG	UNP Q496J9
D	451	VAL	-	EXPRESSION TAG	UNP Q496J9
D	452	PRO	-	EXPRESSION TAG	UNP Q496J9
D	453	ARG	-	EXPRESSION TAG	UNP Q496J9
D	454	GLY	-	EXPRESSION TAG	UNP Q496J9
D	455	SER	-	EXPRESSION TAG	UNP Q496J9

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	6	Total Cl 6 6	0	0

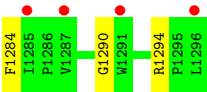
- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Na 2 2	0	0

- Molecule 5 is water.

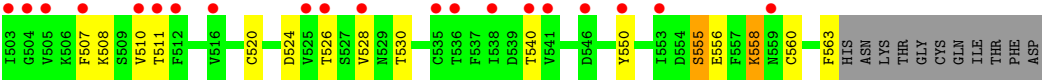
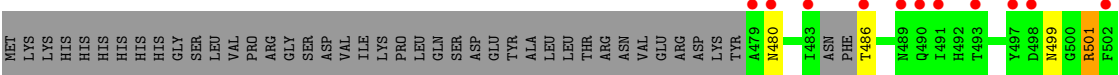
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	50	Total O 50 50	0	0
5	B	18	Total O 18 18	0	0
5	C	6	Total O 6 6	0	0
5	D	8	Total O 8 8	0	0





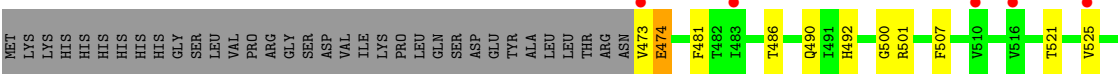
• Molecule 2: Synaptic vesicle glycoprotein 2C

Chain C:



• Molecule 2: Synaptic vesicle glycoprotein 2C

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.44Å 105.26Å 127.96Å 90.00° 90.02° 90.00°	Depositor
Resolution (Å)	19.93 – 2.30 19.93 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.93-2.30) 99.9 (19.93-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 2.30Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1389)	Depositor
R, $R_{free}$	0.235 , 0.269 0.238 , 0.271	Depositor DCC
$R_{free}$ test set	2035 reflections (3.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.6	Xtriage
Anisotropy	0.519	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.47 , 33.9	EDS
Estimated twinning fraction	0.430 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	0 of 67843 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	16539	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 19.29% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, CL

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	0.44	0/3524	0.60	0/4761
1	B	0.33	0/3481	0.52	0/4703
2	C	0.31	0/703	0.48	0/945
2	D	0.46	0/800	0.58	0/1077
All	All	0.39	0/8508	0.56	0/11486

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
2	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	563	PHE	Peptide

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the

Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3442	3412	0	36	0
1	B	3411	3367	0	34	0
2	C	686	634	0	13	0
2	D	779	718	0	6	0
3	A	6	0	0	0	0
4	A	2	0	0	0	0
5	A	50	0	0	7	0
5	B	18	0	0	0	0
5	C	6	0	0	1	0
5	D	8	0	0	1	0
All	All	8408	8131	0	87	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

The worst 5 of 87 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1216:ASN:ND2	5:A:1432:HOH:O	2.14	0.78
1:A:1186:VAL:O	5:A:1425:HOH:O	2.07	0.71
1:B:1265:ASN:OD1	1:B:1268:ASN:ND2	2.23	0.71
2:D:565:ASN:OD1	5:D:608:HOH:O	2.09	0.70
1:A:1179:ARG:NH2	5:A:1413:HOH:O	2.25	0.68

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	413/443 (93%)	385 (93%)	28 (7%)	0	100	100
1	B	405/443 (91%)	349 (86%)	49 (12%)	7 (2%)	14	11
2	C	79/136 (58%)	71 (90%)	8 (10%)	0	100	100
2	D	91/136 (67%)	87 (96%)	3 (3%)	1 (1%)	21	21

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	988/1158 (85%)	892 (90%)	88 (9%)	8 (1%)	27	30

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1166	ALA
2	D	474	GLU
1	B	1140	ARG
1	B	892	SER
1	B	980	TYR

### 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	387/408 (95%)	364 (94%)	23 (6%)	28	35
1	B	382/408 (94%)	351 (92%)	31 (8%)	17	20
2	C	79/128 (62%)	74 (94%)	5 (6%)	25	32
2	D	89/128 (70%)	85 (96%)	4 (4%)	38	50
All	All	937/1072 (87%)	874 (93%)	63 (7%)	23	29

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	969	GLU
1	B	1048	HIS
2	C	558	LYS
1	B	977	SER
1	B	997	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	913	GLN
1	A	988	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 8 ligands modelled in this entry, 8 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.

## 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, 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	416/443 (93%)	0.56	33 (7%) 13 19	12, 22, 45, 64	0
1	B	413/443 (93%)	1.44	118 (28%) 1 1	26, 51, 76, 103	0
2	C	83/136 (61%)	1.89	31 (37%) 1 1	26, 56, 73, 82	0
2	D	93/136 (68%)	0.58	6 (6%) 18 26	14, 27, 46, 66	0
All	All	1005/1158 (86%)	1.04	188 (18%) 2 3	12, 37, 71, 103	0

The worst 5 of 188 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	483	ILE	9.9
1	B	1202	VAL	7.9
1	B	875	ASN	6.8
1	B	878	ILE	6.8
2	C	511	THR	6.4

### 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 carbohydrates in this entry.

### 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	A	1301	1/1	0.18	3.15	37,37,37,37	0
3	CL	A	1308	1/1	0.18	0.04	24,24,24,24	0
3	CL	A	1305	1/1	0.15	-0.35	32,32,32,32	0
4	NA	A	1306	1/1	0.14	-0.60	22,22,22,22	0
3	CL	A	1303	1/1	0.11	-1.59	30,30,30,30	0
3	CL	A	1302	1/1	0.10	-1.95	45,45,45,45	0
3	CL	A	1304	1/1	0.11	-2.12	24,24,24,24	0
4	NA	A	1307	1/1	0.12	-3.24	15,15,15,15	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.