



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 4, 2021 – 05:10 pm BST

PDB ID : 6G5G  
Title : Crystal structure of an engineered Botulinum Neurotoxin type B mutant E1191M/S1199Y in complex with human synaptotagmin 2  
Authors : Masuyer, G.; Elliot, M.; Favre-Guilmard, C.; Liu, S.M.; Maignel, J.; Beard, M.; Carre, D.; Kalinichev, M.; Lezmi, S.; Mir, I.; Nicoleau, C.; Palan, S.; Perier, C.; Raban, E.; Dong, M.; Krupp, J.; Stenmark, P.  
Deposited on : 2018-03-29  
Resolution : 2.00 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](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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.1  
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) : 2.23.1

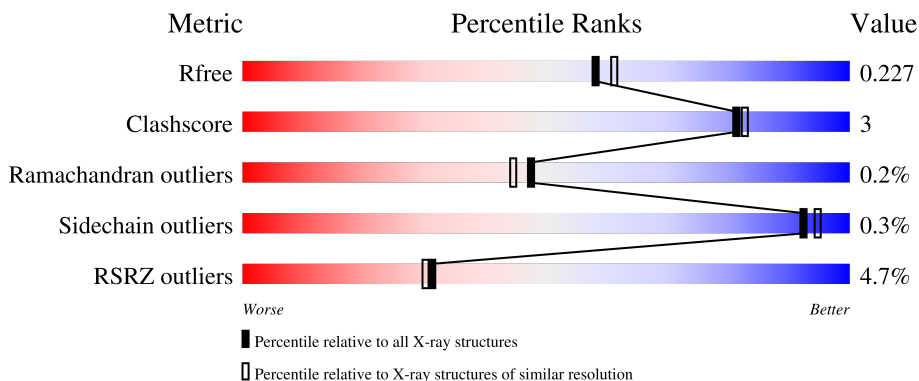
# 1 Overall quality at a glance

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 of 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	1291	
1	B	1291	
2	P	21	

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 11548 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 B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	426	3470	2236	565	654	15	0	0	0
1	B	848	7079	4574	1125	1361	19	0	3	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	231	GLN	GLU	engineered mutation	UNP P10844
A	234	TYR	HIS	engineered mutation	UNP P10844
A	1191	MET	GLU	conflict	UNP P10844
A	1199	TYR	SER	conflict	UNP P10844
B	231	GLN	GLU	engineered mutation	UNP P10844
B	234	TYR	HIS	engineered mutation	UNP P10844
B	1191	MET	GLU	conflict	UNP P10844
B	1199	TYR	SER	conflict	UNP P10844

- Molecule 2 is a protein called Synaptotagmin-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	P	18	156	100	25	30	1	0	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		

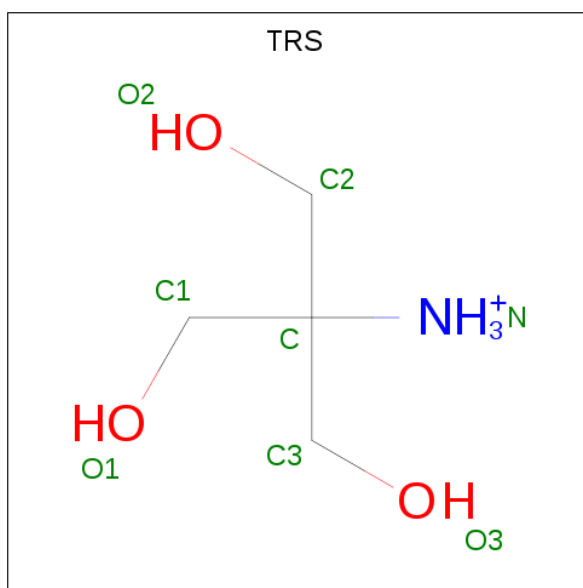
- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			7	4	3		

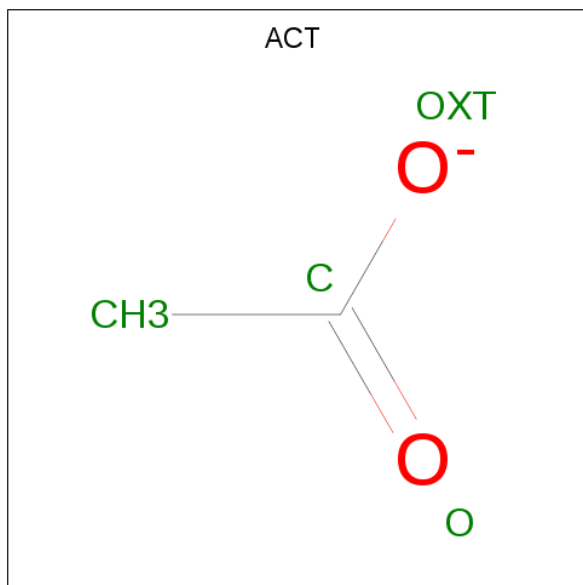
- Molecule 6 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code:

TRS) (formula:  $C_4H_{12}NO_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
6	B	1	8	4	1	3	0	0

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
7	B	1	4	2	2	0	0

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	C	O	0	0
			6	3	3		
8	B	1	Total	C	O	0	0
			6	3	3		
8	B	1	Total	C	O	0	0
			6	3	3		
8	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	173	Total	O	0	0
			173	173		
9	B	601	Total	O	0	0
			601	601		
9	P	14	Total	O	0	0
			14	14		







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	166.60Å 211.05Å 120.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	57.50 – 2.00 54.65 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (57.50-2.00) 99.9 (54.65-2.00)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.65 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0218	Depositor
R, $R_{free}$	0.190 , 0.223 0.197 , 0.227	Depositor DCC
$R_{free}$ test set	7218 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.1	Xtrriage
Anisotropy	0.463	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11548	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.92% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: TRS, PEG, ACT, EDO, MG, GOL

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	0/3548	0.62	0/4787
1	B	0.50	0/7232	0.64	1/9774 (0.0%)
2	P	0.50	0/157	0.52	0/204
All	All	0.48	0/10937	0.63	1/14765 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	933	ARG	NE-CZ-NH1	5.35	122.97	120.30

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	3470	0	3440	14	0
1	B	7079	0	6953	42	0
2	P	156	0	158	0	0
3	A	1	0	0	0	0
4	A	4	0	6	0	0
5	B	14	0	20	2	0
6	B	8	0	12	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	4	0	3	1	0
8	B	24	0	32	1	0
9	A	173	0	0	3	0
9	B	601	0	0	8	0
9	P	14	0	0	0	0
All	All	11548	0	10624	56	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 3.

The worst 5 of 56 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:B:659:LEU:H	1:B:703:GLN:HE22	1.25	0.85
1:B:1222:GLN:HE21	1:B:1279:ASN:HD22	1.34	0.76
1:B:992[A]:ARG:NH2	9:B:1402:HOH:O	2.11	0.75
1:B:910:ARG:NE	9:B:1401:HOH:O	1.98	0.70
1:A:410:ARG:HG2	1:A:416:ILE:HD13	1.73	0.70

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	422/1291 (33%)	411 (97%)	11 (3%)	0	<b>100</b> <b>100</b>
1	B	849/1291 (66%)	820 (97%)	27 (3%)	2 (0%)	47 44
2	P	16/21 (76%)	16 (100%)	0	0	<b>100</b> <b>100</b>
All	All	1287/2603 (49%)	1247 (97%)	38 (3%)	2 (0%)	47 44

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1156	ILE
1	B	922	VAL

### 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	387/1190 (32%)	385 (100%)	2 (0%)	88	92
1	B	790/1190 (66%)	789 (100%)	1 (0%)	93	95
2	P	17/19 (90%)	17 (100%)	0	100	100
All	All	1194/2399 (50%)	1191 (100%)	3 (0%)	92	95

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	8	PHE
1	A	346	ASP
1	B	612	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	943	GLN
1	B	1070	ASN
1	B	958	ASN
1	B	1090	ASN
1	B	757	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](#)

Of 10 ligands modelled in this entry, 1 is monoatomic - leaving 9 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	PEG	B	1302	-	6,6,6	0.62	0	5,5,5	0.39	0
8	GOL	B	1305	-	5,5,5	0.29	0	5,5,5	0.28	0
4	EDO	A	1302	-	3,3,3	0.53	0	2,2,2	0.39	0
8	GOL	B	1308	-	5,5,5	0.26	0	5,5,5	0.46	0
7	ACT	B	1304	-	1,3,3	1.79	0	0,3,3	0.00	-
6	TRS	B	1303	-	7,7,7	0.20	0	9,9,9	0.83	0
8	GOL	B	1306	-	5,5,5	0.30	0	5,5,5	0.26	0
8	GOL	B	1307	-	5,5,5	0.41	0	5,5,5	0.50	0
5	PEG	B	1301	-	6,6,6	0.46	0	5,5,5	0.21	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PEG	B	1302	-	-	2/4/4/4	-
8	GOL	B	1305	-	-	2/4/4/4	-
4	EDO	A	1302	-	-	1/1/1/1	-
8	GOL	B	1308	-	-	0/4/4/4	-
6	TRS	B	1303	-	-	0/9/9/9	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	GOL	B	1306	-	-	2/4/4/4	-
8	GOL	B	1307	-	-	2/4/4/4	-
5	PEG	B	1301	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	B	1306	GOL	C1-C2-C3-O3
8	B	1306	GOL	O2-C2-C3-O3
5	B	1302	PEG	O1-C1-C2-O2
5	B	1302	PEG	O2-C3-C4-O4
8	B	1305	GOL	C1-C2-C3-O3

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1302	PEG	2	0
7	B	1304	ACT	1	0
6	B	1303	TRS	1	0
8	B	1307	GOL	1	0

## 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, 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	426/1291 (32%)	-0.01	10 (2%) 60 59	30, 47, 81, 105	0
1	B	848/1291 (65%)	0.24	51 (6%) 21 20	20, 38, 83, 115	0
2	P	18/21 (85%)	-0.12	0 100 100	30, 44, 92, 95	0
All	All	1292/2603 (49%)	0.15	61 (4%) 31 30	20, 42, 83, 115	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	611	MET	7.6
1	B	819	GLY	5.8
1	B	818	ILE	5.6
1	B	922	VAL	5.3
1	B	815	LEU	4.5

### 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](#)

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. 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	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	ACT	B	1304	4/4	0.73	0.18	57,65,65,68	0
5	PEG	B	1302	7/7	0.80	0.18	45,55,58,59	0
8	GOL	B	1306	6/6	0.86	0.18	59,64,65,68	0
8	GOL	B	1307	6/6	0.87	0.18	51,56,57,59	0
5	PEG	B	1301	7/7	0.89	0.13	67,72,74,75	0
8	GOL	B	1308	6/6	0.90	0.18	70,73,74,74	0
8	GOL	B	1305	6/6	0.91	0.17	41,50,57,59	0
4	EDO	A	1302	4/4	0.93	0.15	43,52,58,63	0
6	TRS	B	1303	8/8	0.96	0.11	40,46,48,50	0
3	MG	A	1301	1/1	0.97	0.05	46,46,46,46	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.