



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 1, 2022 – 04:22 pm BST

PDB ID : 6G5F
Title : Crystal structure of an engineered Botulinum Neurotoxin type B mutant E1191M/S1199Y in complex with human synaptotagmin 1
Authors : Masuyer, G.; Elliot, M.; Favre-Guilnard, 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 : 2.50 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 types of validation reports are described at <http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.30
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.30

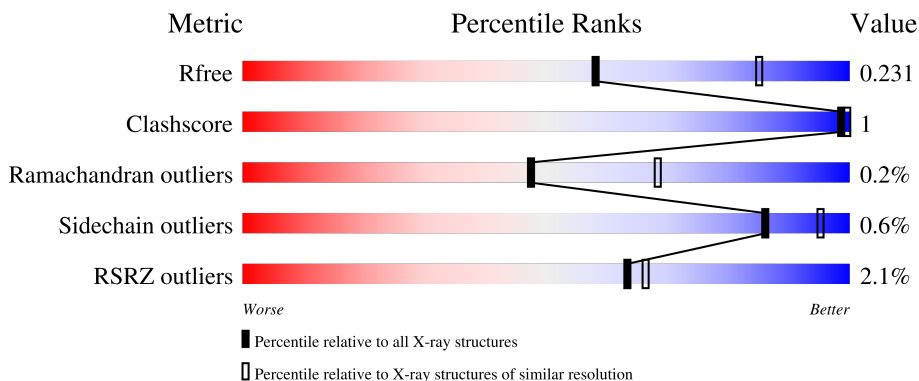
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 $\leq 5\%$. 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 5 unique types of molecules in this entry. The entry contains 11000 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	440	3592	2307	590	680	15	0	5	0
1	B	845	6959	4497	1103	1340	19	0	0	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	engineered mutation	UNP P10844
A	1199	TYR	SER	engineered mutation	UNP P10844
B	231	GLN	GLU	engineered mutation	UNP P10844
B	234	TYR	HIS	engineered mutation	UNP P10844
B	1191	MET	GLU	engineered mutation	UNP P10844
B	1199	TYR	SER	engineered mutation	UNP P10844

- Molecule 2 is a protein called Synaptotagmin-1.

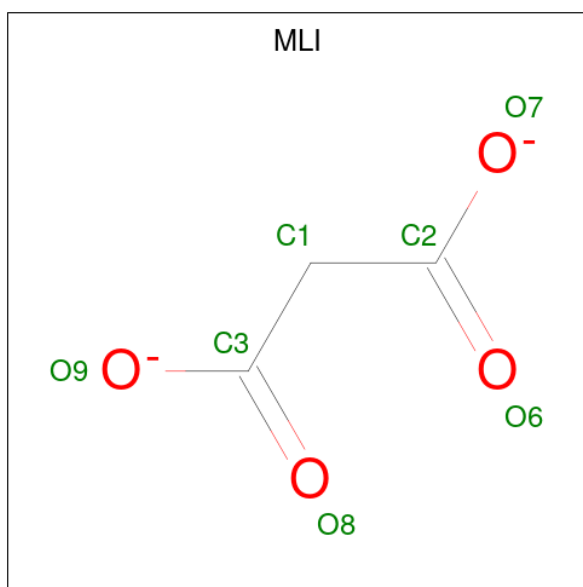
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	P	18	156	100	26	29	1	0	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 4 is MALONATE ION (three-letter code: MLI) (formula: $C_3H_2O_4$).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	153	Total 153	O 153	0	0
5	B	115	Total 115	O 115	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	176.08Å 356.42Å 103.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	98.50 – 2.50 98.48 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (98.50-2.50) 100.0 (98.48-2.50)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.80 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.202 , 0.232 0.204 , 0.231	Depositor DCC
R_{free} test set	5577 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	54.6	Xtriage
Anisotropy	0.601	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11000	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MLI, 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	0.43	0/3672	0.62	0/4956
1	B	0.43	0/7108	0.58	0/9616
2	P	0.44	0/158	0.46	0/204
All	All	0.43	0/10938	0.59	0/14776

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	3592	0	3546	3	0
1	B	6959	0	6780	8	0
2	P	156	0	158	0	0
3	A	18	0	24	1	0
4	B	7	0	2	0	0
5	A	153	0	0	1	0
5	B	115	0	0	1	0
All	All	11000	0	10510	11	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 1.

The worst 5 of 11 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:213[B]:SER:O	1:A:214[B]:ILE:CB	2.33	0.76
1:B:704:TRP:CD1	1:B:787:MET:HE1	2.50	0.47
1:B:977:ASP:HB3	1:B:1031:ILE:HG23	1.98	0.46
1:B:845:ASN:ND2	5:B:1402:HOH:O	2.51	0.43
1:B:1008:THR:HG23	1:B:1065:GLU:HG3	2.01	0.43

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	443/1291 (34%)	426 (96%)	15 (3%)	2 (0%)	29	48
1	B	839/1291 (65%)	805 (96%)	32 (4%)	2 (0%)	47	68
2	P	16/21 (76%)	16 (100%)	0	0	100	100
All	All	1298/2603 (50%)	1247 (96%)	47 (4%)	4 (0%)	47	61

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	214[A]	ILE
1	A	214[B]	ILE
1	B	1132	LYS
1	B	821	ALA

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	397/1190 (33%)	394 (99%)	3 (1%)	81	93
1	B	767/1190 (64%)	763 (100%)	4 (0%)	88	96
2	P	17/18 (94%)	17 (100%)	0	100	100
All	All	1181/2398 (49%)	1174 (99%)	7 (1%)	86	95

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	641	ASN
1	B	687	THR
1	B	1008	THR
1	B	975	LEU
1	A	440	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1105	ASN
1	B	1107	ASN
1	B	1176	GLN
1	B	1173	ASN
1	B	873	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

4 ligands are modelled in this entry.

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
3	GOL	A	1302	-	5,5,5	0.46	0	5,5,5	0.36	0
3	GOL	A	1301	-	5,5,5	0.39	0	5,5,5	0.31	0
4	MLI	B	1301	-	6,6,6	1.10	0	7,7,7	1.03	0
3	GOL	A	1303	-	5,5,5	0.30	0	5,5,5	0.31	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
3	GOL	A	1302	-	-	0/4/4/4	-
3	GOL	A	1301	-	-	2/4/4/4	-
4	MLI	B	1301	-	-	2/4/4/4	-
3	GOL	A	1303	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1301	GOL	C1-C2-C3-O3
3	A	1301	GOL	O2-C2-C3-O3
3	A	1303	GOL	O1-C1-C2-O2
3	A	1303	GOL	O1-C1-C2-C3
4	B	1301	MLI	C3-C1-C2-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1303	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, 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	440/1291 (34%)	0.01	2 (0%) 91 91	41, 55, 92, 131	0
1	B	845/1291 (65%)	0.16	19 (2%) 62 65	45, 81, 123, 167	0
2	P	18/21 (85%)	1.32	7 (38%) 0 0	109, 131, 161, 163	0
All	All	1303/2603 (50%)	0.13	28 (2%) 63 66	41, 70, 122, 167	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	40	PHE	4.0
2	P	47	PHE	3.4
1	B	945	TYR	3.3
1	A	217	ARG	3.2
2	P	38	ASP	3.2

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, 95th 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(\AA^2)	Q<0.9
4	MLI	B	1301	7/7	0.84	0.31	76,84,89,93	0
3	GOL	A	1302	6/6	0.85	0.20	78,80,81,81	0
3	GOL	A	1303	6/6	0.88	0.23	66,74,77,82	0
3	GOL	A	1301	6/6	0.88	0.18	84,88,91,93	0

6.5 Other polymers [i](#)

There are no such residues in this entry.