



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 05:45 am BST

PDB ID : 6B4S
Title : Crystal Structure of Brazil nut (*Bertholletia excelsa*) allergen Ber e 2
Authors : Zhang, Y.Z.; Guo, F.
Deposited on : 2017-09-27
Resolution : 2.04 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 : 2.11
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) : 2.11

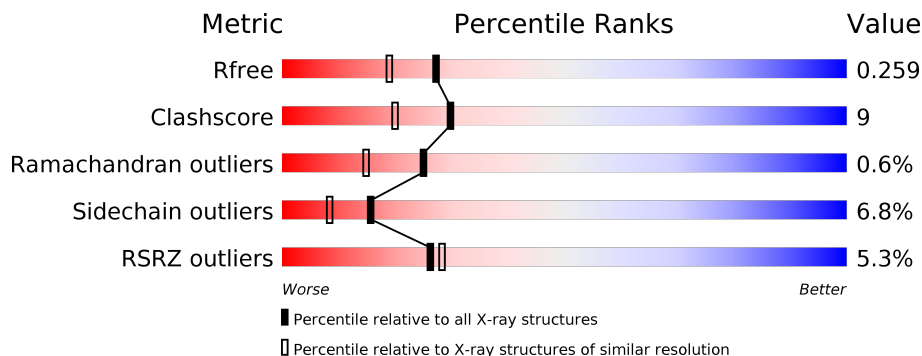
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.04 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	426	
1	B	426	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 11052 atoms, of which 5434 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 11S globulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	364	5588	1777	2762	514	524	11	0	0	0
1	B	351	5388	1720	2672	490	495	11	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	454	ASN	GLN	conflict	UNP Q84ND2
B	454	ASN	GLN	conflict	UNP Q84ND2

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	37	Total	O	0	0
			37	37		
2	B	39	Total	O	0	0
			39	39		

4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	92.43Å 92.43Å 213.19Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.21 – 2.04 46.21 – 2.04	Depositor EDS
% Data completeness (in resolution range)	99.1 (46.21-2.04) 93.4 (46.21-2.04)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.20 (at 2.03Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.250 , 0.270 0.245 , 0.259	Depositor DCC
R_{free} test set	2016 reflections (4.67%)	wwPDB-VP
Wilson B-factor (Å ²)	42.7	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 49.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	0.236 for h,-h-k,-l	Xtriage
Reported twinning fraction	0.260 for h,-h-k,-l	Depositor
Outliers	0 of 43199 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	11052	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.14% 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

5.1 Standard geometry

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	0.28	0/2876	0.49	0/3904
1	B	0.29	0/2765	0.53	2/3754 (0.1%)
All	All	0.29	0/5641	0.51	2/7658 (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
1	B	1	0
All	All	1	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	220	PHE	CB-CA-C	5.81	122.01	110.40
1	B	220	PHE	N-CA-CB	5.07	119.72	110.60

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	220	PHE	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	259	LYS	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	2826	2762	2762	58	1
1	B	2716	2672	2672	49	0
2	A	37	0	0	12	0
2	B	39	0	0	9	0
All	All	5618	5434	5434	103	1

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 9.

The worst 5 of 103 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:53:GLU:O	2:A:501:HOH:O	1.81	0.96
1:B:59:ASP:OD2	2:B:501:HOH:O	1.98	0.81
1:A:190:GLY:N	1:A:215:ASN:OD1	2.21	0.73
1:A:255:GLU:OE1	1:A:256:GLN:N	2.20	0.73
1:B:37:THR:OG1	1:B:61:GLN:OE1	2.08	0.71

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:219:GLY:O	1:A:369:THR:OG1[2_665]	2.11	0.09

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	354/426 (83%)	324 (92%)	26 (7%)	4 (1%)	14	5
1	B	337/426 (79%)	315 (94%)	22 (6%)	0	100	100
All	All	691/852 (81%)	639 (92%)	48 (7%)	4 (1%)	25	15

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	236	GLU
1	A	240	LYS
1	A	245	ASP
1	A	239	ARG

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	294/355 (83%)	277 (94%)	17 (6%)	20	11
1	B	283/355 (80%)	261 (92%)	22 (8%)	12	5
All	All	577/710 (81%)	538 (93%)	39 (7%)	16	8

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	30	GLU
1	B	104	MET
1	B	412	SER
1	B	35	ARG
1	B	44	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates 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	364/426 (85%)	0.52	13 (3%) 42 46	32, 48, 64, 80	0
1	B	351/426 (82%)	0.78	25 (7%) 16 17	33, 54, 87, 164	0
All	All	715/852 (83%)	0.65	38 (5%) 26 28	32, 51, 74, 164	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	243	GLY	8.5
1	A	258	LEU	7.5
1	B	263	PRO	7.5
1	B	212	LEU	6.3
1	B	216	VAL	4.6

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.