



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 02:01 pm GMT

PDB ID : 2BTO  
Title : STRUCTURE OF BTUBA FROM PROSTHECOBACTER DEJONGEII  
Authors : Schlieper, D.; Lowe, J.  
Deposited on : 2005-06-04  
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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

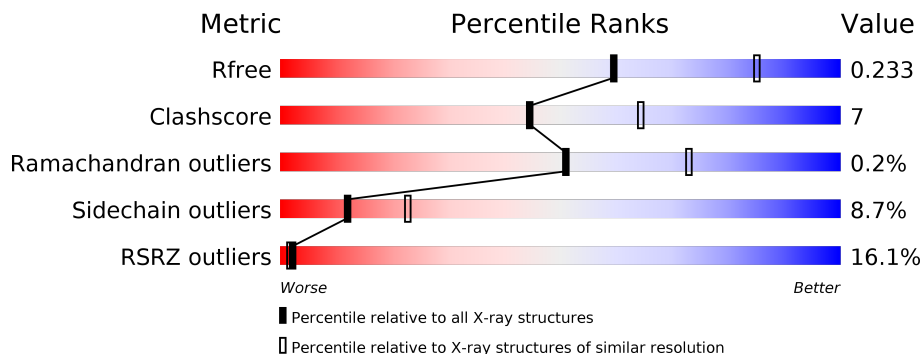
# 1 Overall quality at a glance

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))
$R_{free}$	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	473	<div> <div>11%</div> <div> <div></div> <div>70%</div> <div>16%</div> <div>•</div> <div>13%</div> </div> </div>
1	B	473	<div> <div>15%</div> <div> <div></div> <div>70%</div> <div>17%</div> <div>•</div> <div>11%</div> </div> </div>
2	T	108	<div> <div>26%</div> <div> <div></div> <div>75%</div> <div>16%</div> <div>5%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7451 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 TUBULIN BTUBA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	413	Total	C	N	O	S	0	0	0
			3156	2012	531	597	16			
1	B	423	Total	C	N	O	S	0	0	0
			3230	2055	544	614	17			

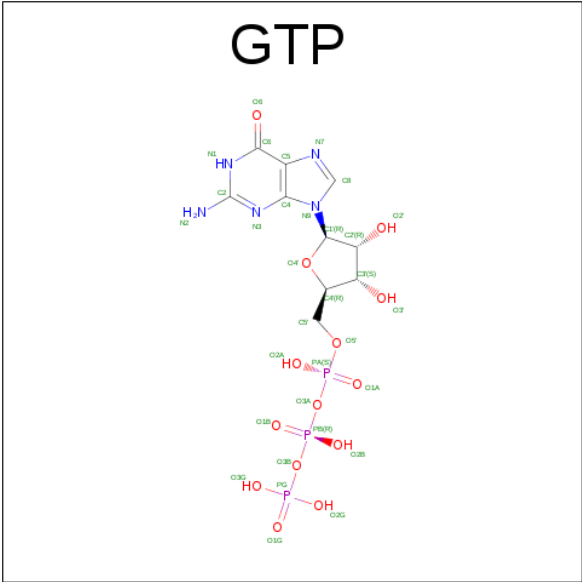
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	255	SER	THR	CONFLICT	UNP Q8GCC5
B	255	SER	THR	CONFLICT	UNP Q8GCC5

- Molecule 2 is a protein called THIOREDOXIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	103	Total	C	N	O	S	0	0	0
			785	506	126	150	3			

- Molecule 3 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>14</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	B	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	108	Total	O	0	0
			108	108		
4	B	93	Total	O	0	0
			93	93		
4	T	15	Total	O	0	0
			15	15		



- Molecule 1: TUBULIN BTUBA



Chain T:

Sequence logo for Chain T. The y-axis represents frequency in bits (0.00 to 1.50). The x-axis shows positions 1 to 200. A bar chart at the top indicates the percentage of residues in each category: 26% (red), 75% (green), 16% (yellow), 5% (orange), and 5% (grey).

Key residues (positions and amino acids) are highlighted:

- Position 1: SER
- Position 2: ASP
- Position 3: LYS
- Position 4: I23
- Position 5: I24
- Position 6: D28
- Position 7: D29
- Position 8: T33
- Position 9: D34
- Position 10: K37
- Position 11: A38
- Position 12: D39
- Position 13: G40
- Position 14: A41
- Position 15: L42
- Position 16: L43
- Position 17: F46
- Position 18: V47
- Position 19: A48
- Position 20: E49
- Position 21: W50
- Position 22: C51
- Position 23: A65
- Position 24: D66
- Position 25: E67
- Position 26: Y68
- Position 27: Q69
- Position 28: G70
- Position 29: K71
- Position 30: I72
- Position 31: T73
- Position 32: V74
- Position 33: N78
- Position 34: I79
- Position 35: D80
- Position 36: A81
- Position 37: A86
- Position 38: I91
- Position 39: I94
- Position 40: L97
- Position 41: L98
- Position 42: L99
- Position 43: F100
- Position 44: K101
- Position 45: M102
- Position 46: G103
- Position 47: E104
- Position 48: K109
- Position 49: L113
- Position 50: S114
- Position 51: K115
- Position 52: G116
- Position 53: Q117
- Position 54: L118
- Position 55: K119
- Position 56: E120
- Position 57: F121
- Position 58: A124
- Position 59: M125
- Position 60: LEU
- Position 61: ALA

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	180.54Å 180.54Å 84.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.50 31.62 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-2.50) 99.9 (31.62-2.50)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.79 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.200 , 0.235 0.198 , 0.233	Depositor DCC
$R_{free}$ test set	2733 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	44.5	Xtriage
Anisotropy	0.449	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 66.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7451	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% 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

### 5.1 Standard geometry

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

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	0.50	0/3219	0.64	0/4372
1	B	1.26	6/3294 (0.2%)	0.66	2/4471 (0.0%)
2	T	0.68	2/800 (0.2%)	0.54	0/1086
All	All	0.94	8/7313 (0.1%)	0.64	2/9929 (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	B	0	2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	410	GLU	CD-OE1	50.64	1.81	1.25
1	B	411	GLU	CD-OE2	26.13	1.54	1.25
1	B	411	GLU	CD-OE1	23.18	1.51	1.25
1	B	411	GLU	CB-CG	18.88	1.88	1.52
2	T	67	GLU	CD-OE2	13.83	1.40	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	411	GLU	OE1-CD-OE2	-11.20	109.86	123.30
1	B	410	GLU	CG-CD-OE2	-5.63	107.05	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	B	410	GLU	Sidechain
1	B	411	GLU	Sidechain

## 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	3156	0	3143	44	0
1	B	3230	0	3221	55	0
2	T	785	0	790	13	0
3	A	32	0	12	3	0
3	B	32	0	12	1	0
4	A	108	0	0	10	0
4	B	93	0	0	5	0
4	T	15	0	0	2	0
All	All	7451	0	7178	108	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 7.

The worst 5 of 108 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:411:GLU:CG	1:B:411:GLU:CB	1.88	1.51
1:A:165:GLU:HG2	4:A:2055:HOH:O	1.36	1.21
1:B:410:GLU:CD	1:B:410:GLU:OE1	1.81	1.16
3:A:1433:GTP:PB	4:A:2106:HOH:O	2.22	0.97
1:B:235:GLU:OE1	1:B:366:HIS:HE1	1.67	0.77

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	403/473 (85%)	389 (96%)	14 (4%)	0	100	100
1	B	417/473 (88%)	399 (96%)	16 (4%)	2 (0%)	32	53
2	T	101/108 (94%)	99 (98%)	2 (2%)	0	100	100
All	All	921/1054 (87%)	887 (96%)	32 (4%)	2 (0%)	51	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	164	GLY
1	B	398	ALA

### 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	341/387 (88%)	312 (92%)	29 (8%)	12	23
1	B	350/387 (90%)	321 (92%)	29 (8%)	13	25
2	T	83/87 (95%)	74 (89%)	9 (11%)	7	14
All	All	774/861 (90%)	707 (91%)	67 (9%)	12	23

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	63	SER
1	B	204	LEU
2	T	91	ILE
1	B	88	LEU
1	B	140	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	266	GLN
1	B	361	GLN
1	B	405	ASN
1	A	412	GLN
1	B	414	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 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	GTP	A	1433	-	27,34,34	1.12	2 (7%)	27,54,54	1.94	6 (22%)
3	GTP	B	1433	-	27,34,34	1.06	1 (3%)	27,54,54	1.76	5 (18%)

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	GTP	A	1433	-	-	0/18/38/38	0/3/3/3
3	GTP	B	1433	-	-	0/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1433	GTP	PG-O3B	2.03	1.63	1.60
3	B	1433	GTP	C6-N1	3.30	1.39	1.33
3	A	1433	GTP	C6-N1	3.45	1.39	1.33

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1433	GTP	N3-C2-N1	-5.23	119.81	127.46
3	B	1433	GTP	N3-C2-N1	-5.09	120.03	127.46
3	B	1433	GTP	C5-C6-N1	-3.14	119.02	123.48
3	A	1433	GTP	C5-C6-N1	-3.12	119.04	123.48
3	A	1433	GTP	C1'-N9-C4	-2.56	122.22	126.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1433	GTP	3	0
3	B	1433	GTP	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	413/473 (87%)	0.62	51 (12%)	5 4	42, 51, 63, 76	0
1	B	423/473 (89%)	0.76	72 (17%)	2 1	39, 51, 66, 78	0
2	T	103/108 (95%)	1.37	28 (27%)	1 0	66, 71, 77, 78	0
All	All	939/1054 (89%)	0.77	151 (16%)	2 2	39, 52, 73, 78	0

The worst 5 of 151 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	398	ALA	8.2
1	B	395	GLN	6.6
1	A	41	PRO	6.4
2	T	40	GLY	6.0
1	B	411	GLU	5.8

### 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](#)

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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GTP	A	1433	32/32	0.95	0.18	-0.01	48,54,57,58	0
3	GTP	B	1433	32/32	0.90	0.15	-0.69	70,72,75,76	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.