



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 13, 2024 – 09:53 AM EDT

PDB ID : 1N4D
Title : The Ligand-Free Structure of E coli BtuF, the Periplasmic Binding Protein for Vitamin B12
Authors : Karpowich, N.; Smith, P.C.; Hunt, J.F.
Deposited on : 2002-10-30
Resolution : 3.00 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

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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36.2
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) : 2.36.2

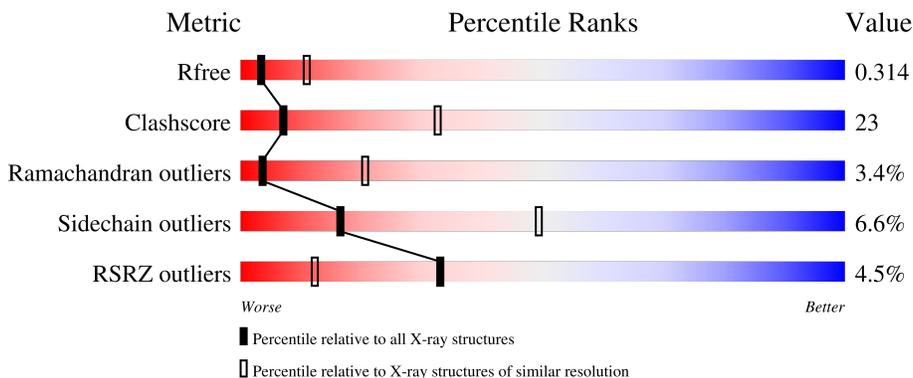
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 ≥ 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	252	
1	B	252	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 3661 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 Vitamin B12 transport protein btuF.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	244	1902	1213	331	354	2	2	0	0	0
1	B	227	1759	1119	307	329	2	2	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	MSE	MET	MODIFIED RESIDUE	UNP P37028
A	83	MSE	MET	MODIFIED RESIDUE	UNP P37028
A	245	LEU	-	EXPRESSION TAG	UNP P37028
A	246	GLU	-	EXPRESSION TAG	UNP P37028
A	247	HIS	-	EXPRESSION TAG	UNP P37028
A	248	HIS	-	EXPRESSION TAG	UNP P37028
A	249	HIS	-	EXPRESSION TAG	UNP P37028
A	250	HIS	-	EXPRESSION TAG	UNP P37028
A	251	HIS	-	EXPRESSION TAG	UNP P37028
A	252	HIS	-	EXPRESSION TAG	UNP P37028
B	1047	MSE	MET	MODIFIED RESIDUE	UNP P37028
B	1083	MSE	MET	MODIFIED RESIDUE	UNP P37028
B	1245	LEU	-	EXPRESSION TAG	UNP P37028
B	1246	GLU	-	EXPRESSION TAG	UNP P37028
B	1247	HIS	-	EXPRESSION TAG	UNP P37028
B	1248	HIS	-	EXPRESSION TAG	UNP P37028
B	1249	HIS	-	EXPRESSION TAG	UNP P37028
B	1250	HIS	-	EXPRESSION TAG	UNP P37028
B	1251	HIS	-	EXPRESSION TAG	UNP P37028
B	1252	HIS	-	EXPRESSION TAG	UNP P37028

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	56.25Å 84.08Å 209.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 19.42 – 2.48	Depositor EDS
% Data completeness (in resolution range)	12.7 (20.00-3.00) 66.4 (19.42-2.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.50 (at 2.49Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.242 , 0.298 0.244 , 0.314	Depositor DCC
R_{free} test set	614 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	50.9	Xtrriage
Anisotropy	0.507	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 61.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	3661	wwPDB-VP
Average B, all atoms (Å ²)	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.65% 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](#)

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	0.46	0/1944	0.63	0/2648
1	B	0.41	0/1795	0.67	1/2445 (0.0%)
All	All	0.44	0/3739	0.65	1/5093 (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	1143	ASN	C-N-CD	-9.38	99.96	120.60

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	1902	0	1922	77	0
1	B	1759	0	1775	90	0
All	All	3661	0	3697	167	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 23.

The worst 5 of 167 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:193:THR:HG21	1:A:222:PHE:HD2	1.40	0.86
1:A:123:GLN:HE21	1:A:123:GLN:HA	1.40	0.85
1:A:35:ALA:HA	1:A:38:ILE:HD12	1.59	0.84
1:A:112:GLN:HA	1:A:112:GLN:HE21	1.41	0.83
1:A:165:ASN:HD22	1:A:166:ILE:N	1.75	0.83

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	242/252 (96%)	209 (86%)	26 (11%)	7 (3%)	4	24
1	B	223/252 (88%)	187 (84%)	27 (12%)	9 (4%)	3	17
All	All	465/504 (92%)	396 (85%)	53 (11%)	16 (3%)	3	20

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1113	ALA
1	B	1145	PRO
1	B	1219	SER
1	A	133	LYS
1	A	150	LYS

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	205/211 (97%)	191 (93%)	14 (7%)	16	48
1	B	190/211 (90%)	178 (94%)	12 (6%)	18	51
All	All	395/422 (94%)	369 (93%)	26 (7%)	16	49

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1007	LEU
1	B	1050	GLU
1	B	1146	PHE
1	B	1039	GLU
1	B	1055	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1048	ASN
1	B	1092	GLN
1	B	1074	GLN
1	B	1099	GLN
1	A	139	GLN

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](#)

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	242/252 (96%)	-0.44	9 (3%) 41 17	15, 46, 124, 146	0
1	B	225/252 (89%)	-0.19	12 (5%) 26 10	16, 59, 133, 142	0
All	All	467/504 (92%)	-0.32	21 (4%) 33 12	15, 52, 130, 146	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	209	GLN	5.4
1	B	1132	PRO	4.8
1	B	1126	ALA	3.7
1	A	142	ILE	3.7
1	B	1143	ASN	3.5

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.