



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 01:51 PM UTC

PDB ID : 8BMX / pdb_00008bmx
Title : Bacteroides thetaiotaomicron B12 TonB dependent transporter in complex with a surface lipoprotein
Authors : Abellon-Ruiz, J.; Jana, K.; Silale, A.; Basle, A.; Kleinekathofer, U.; van den Berg, B.
Deposited on : 2022-11-11
Resolution : 3.72 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

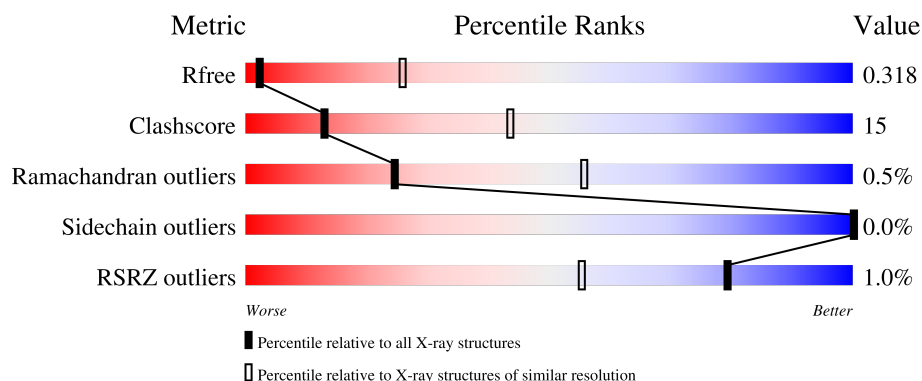
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.72 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}	180053	1182 (3.84-3.60)
Clashscore	190562	1222 (3.84-3.60)
Ramachandran outliers	187476	1178 (3.84-3.60)
Sidechain outliers	187428	1174 (3.84-3.60)
RSRZ outliers	180081	1181 (3.84-3.60)

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	698	57% 32% 10%
1	B	698	2% 58% 32% 10%
1	C	698	% 59% 31% 9%
2	D	366	66% 25% 8%
2	E	366	% 68% 23% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	366	% 66% 25% 8%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 23366 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 Putative TonB-linked outer membrane receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	633	Total	C	N	O	S	0	0	0
			5078	3239	860	960	19			
1	A	628	Total	C	N	O	S	0	0	0
			5038	3213	853	953	19			
1	B	627	Total	C	N	O	S	0	0	0
			5021	3200	851	951	19			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	693	HIS	-	expression tag	UNP Q8A6D1
C	694	HIS	-	expression tag	UNP Q8A6D1
C	695	HIS	-	expression tag	UNP Q8A6D1
C	696	HIS	-	expression tag	UNP Q8A6D1
C	697	HIS	-	expression tag	UNP Q8A6D1
C	698	HIS	-	expression tag	UNP Q8A6D1
A	693	HIS	-	expression tag	UNP Q8A6D1
A	694	HIS	-	expression tag	UNP Q8A6D1
A	695	HIS	-	expression tag	UNP Q8A6D1
A	696	HIS	-	expression tag	UNP Q8A6D1
A	697	HIS	-	expression tag	UNP Q8A6D1
A	698	HIS	-	expression tag	UNP Q8A6D1
B	693	HIS	-	expression tag	UNP Q8A6D1
B	694	HIS	-	expression tag	UNP Q8A6D1
B	695	HIS	-	expression tag	UNP Q8A6D1
B	696	HIS	-	expression tag	UNP Q8A6D1
B	697	HIS	-	expression tag	UNP Q8A6D1
B	698	HIS	-	expression tag	UNP Q8A6D1

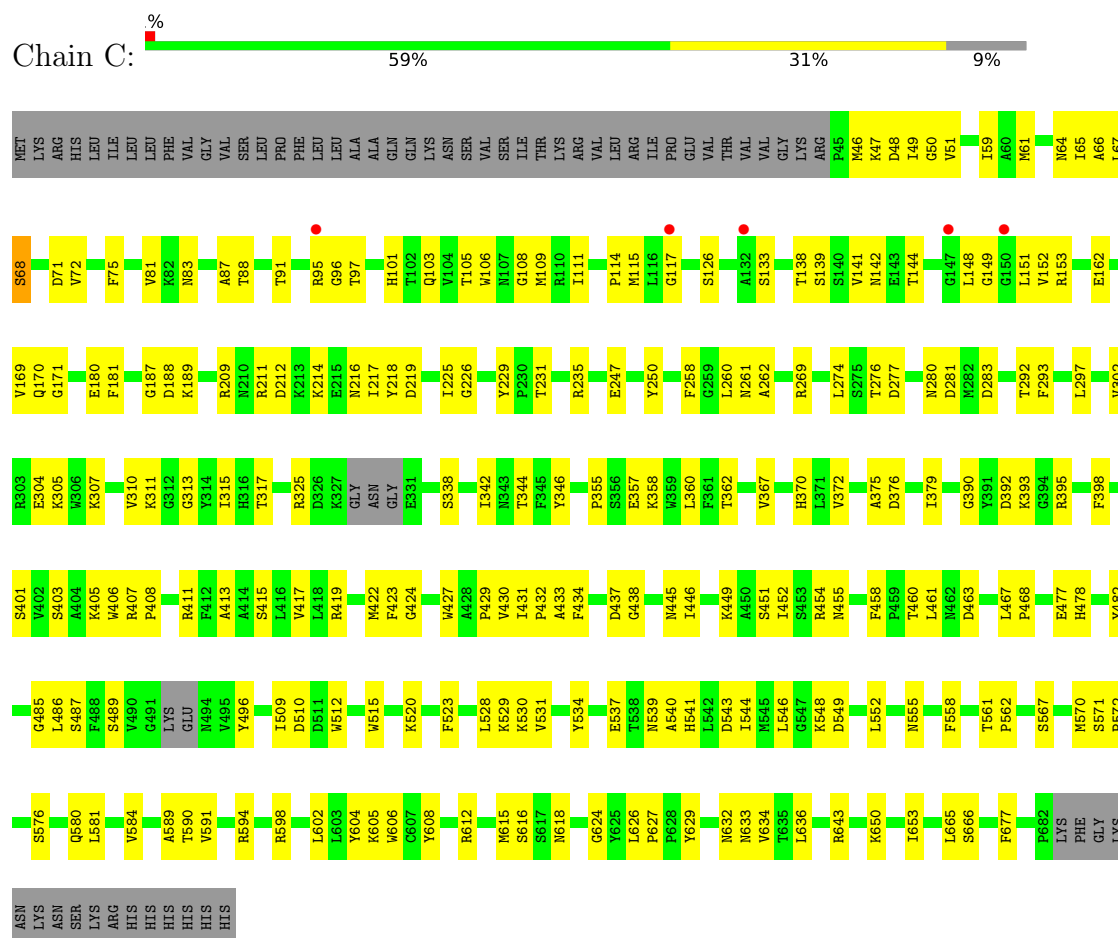
- Molecule 2 is a protein called YncE family protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	335	Total 2743	C 1769	N 434	O 526	S 14	0	0	0
2	E	335	Total 2743	C 1769	N 434	O 526	S 14	0	0	0
2	F	335	Total 2743	C 1769	N 434	O 526	S 14	0	0	0

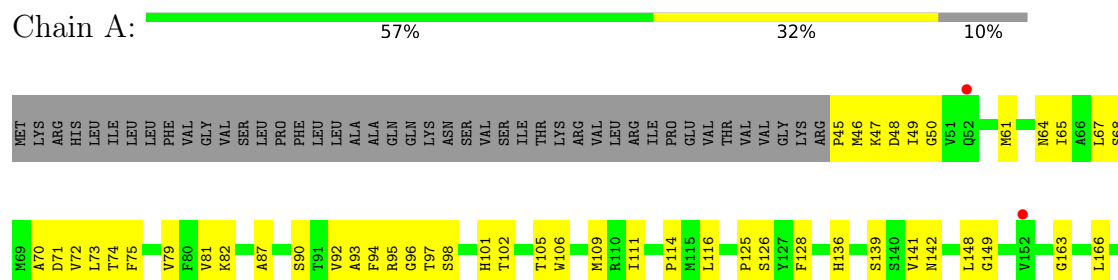
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Putative TonB-linked outer membrane receptor



- Molecule 1: Putative TonB-linked outer membrane receptor



- Molecule 2: YncE family protein

Chain D: 



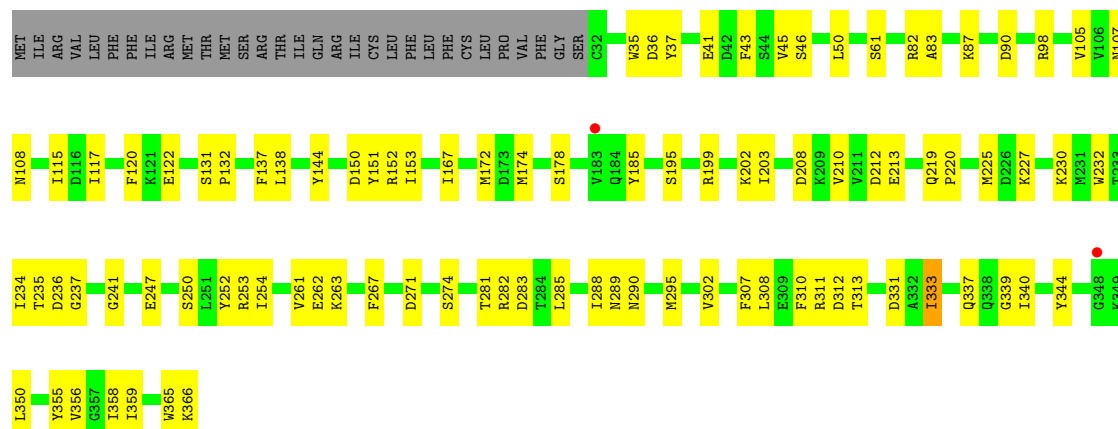
- Molecule 2: YncE family protein

Chain E: 



- Molecule 2: YncE family protein

Chain F: 



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	280.25Å 154.30Å 107.51Å 90.00° 106.96° 90.00°	Depositor
Resolution (Å)	72.06 – 3.72 72.06 – 3.72	Depositor EDS
% Data completeness (in resolution range)	99.1 (72.06-3.72) 99.2 (72.06-3.72)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.32 (at 3.67Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.253 , 0.318 0.253 , 0.318	Depositor DCC
R_{free} test set	2232 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å ²)	110.2	Xtriage
Anisotropy	0.258	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 116.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	23366	wwPDB-VP
Average B, all atoms (Å ²)	124.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% 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	0.18	0/5176	0.47	0/7013
1	B	0.16	0/5158	0.45	1/6989 (0.0%)
1	C	0.16	0/5218	0.46	0/7071
2	D	0.22	0/2820	0.73	0/3836
2	E	0.20	0/2820	0.73	3/3836 (0.1%)
2	F	0.19	0/2820	0.72	3/3836 (0.1%)
All	All	0.18	0/24012	0.57	7/32581 (0.0%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	290	ASN	CA-C-N	6.25	132.04	122.11
2	F	290	ASN	C-N-CA	6.25	132.04	122.11
2	E	290	ASN	N-CA-C	-6.08	107.08	114.75
2	E	290	ASN	CA-C-N	5.70	131.22	121.87
2	E	290	ASN	C-N-CA	5.70	131.22	121.87

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	5038	0	4827	172	0
1	B	5021	0	4806	172	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	5078	0	4868	180	0
2	D	2743	0	2601	76	0
2	E	2743	0	2601	60	0
2	F	2743	0	2601	70	0
All	All	23366	0	22304	685	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 15.

The worst 5 of 685 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:C:219:ASP:HB3	1:C:225:ILE:HD11	1.44	0.99
1:C:235:ARG:HG2	1:C:281:ASP:HB2	1.56	0.87
1:A:442:LYS:H	1:A:442:LYS:HD2	1.40	0.83
1:C:247:GLU:HB3	1:C:261:ASN:HB3	1.62	0.81
1:B:362:THR:HG1	1:B:403:SER:HG	1.29	0.81

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	620/698 (89%)	575 (93%)	43 (7%)	2 (0%)	36	65
1	B	619/698 (89%)	575 (93%)	41 (7%)	3 (0%)	24	56
1	C	627/698 (90%)	584 (93%)	41 (6%)	2 (0%)	36	65
2	D	333/366 (91%)	313 (94%)	18 (5%)	2 (1%)	21	52
2	E	333/366 (91%)	314 (94%)	17 (5%)	2 (1%)	21	52
2	F	333/366 (91%)	314 (94%)	17 (5%)	2 (1%)	21	52

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2865/3192 (90%)	2675 (93%)	177 (6%)	13 (0%)	24	56

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	68	SER
1	A	68	SER
1	A	306	TRP
1	B	68	SER
1	B	644	TRP

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	545/607 (90%)	545 (100%)	0	100	100
1	B	543/607 (90%)	542 (100%)	1 (0%)	87	86
1	C	549/607 (90%)	549 (100%)	0	100	100
2	D	297/327 (91%)	297 (100%)	0	100	100
2	E	297/327 (91%)	297 (100%)	0	100	100
2	F	297/327 (91%)	297 (100%)	0	100	100
All	All	2528/2802 (90%)	2527 (100%)	1 (0%)	100	100

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	446	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	277	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	290	ASN
2	F	347	GLN
2	F	219	GLN
2	E	118	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 oligosaccharides 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	628/698 (89%)	-0.09	3 (0%) 87 69	77, 113, 166, 271	0
1	B	627/698 (89%)	0.04	15 (2%) 59 37	98, 141, 187, 248	0
1	C	633/698 (90%)	-0.08	5 (0%) 82 60	72, 108, 158, 251	0
2	D	335/366 (91%)	-0.18	0 100 100	81, 109, 153, 214	0
2	E	335/366 (91%)	-0.16	3 (0%) 81 58	88, 114, 153, 191	0
2	F	335/366 (91%)	-0.11	2 (0%) 85 66	91, 135, 180, 260	0
All	All	2893/3192 (90%)	-0.08	28 (0%) 79 55	72, 119, 173, 271	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	130	ASP	4.3
2	F	183	VAL	4.1
1	B	131	ASP	3.7
1	C	95	ARG	3.1
1	B	536	ALA	2.8

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 oligosaccharides 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.