



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 01:11 am BST

PDB ID : 3V9F
Title : Crystal Structure of the extracellular domain of the putative hybrid two component system BT3049 from *B. thetaiotaomicron*
Authors : Zhang, Z.; Liu, Q.; Hendrickson, W.A.
Deposited on : 2011-12-27
Resolution : 3.30 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

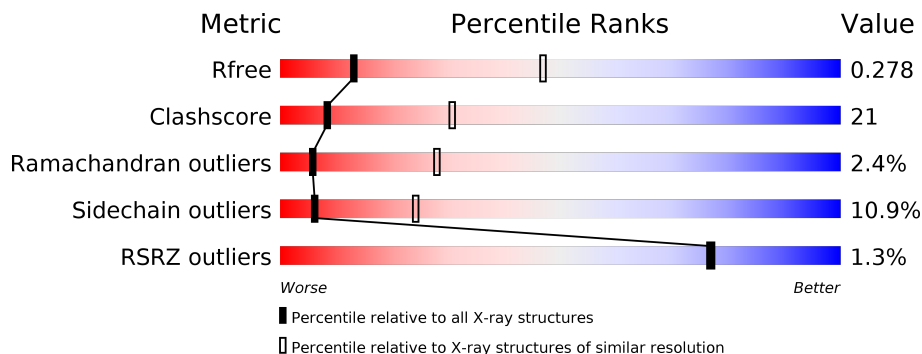
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	781	
1	B	781	
1	C	781	
1	D	781	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 23560 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 Two-component system sensor histidine kinase/response regulator, hybrid (One-component system).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	S				Se
1	A	736	5890	3751	990	1134	10	5	0	0	0
1	B	736	5890	3751	990	1134	10	5	0	0	0
1	C	736	5890	3751	990	1134	10	5	0	0	0
1	D	736	5890	3751	990	1134	10	5	0	0	0

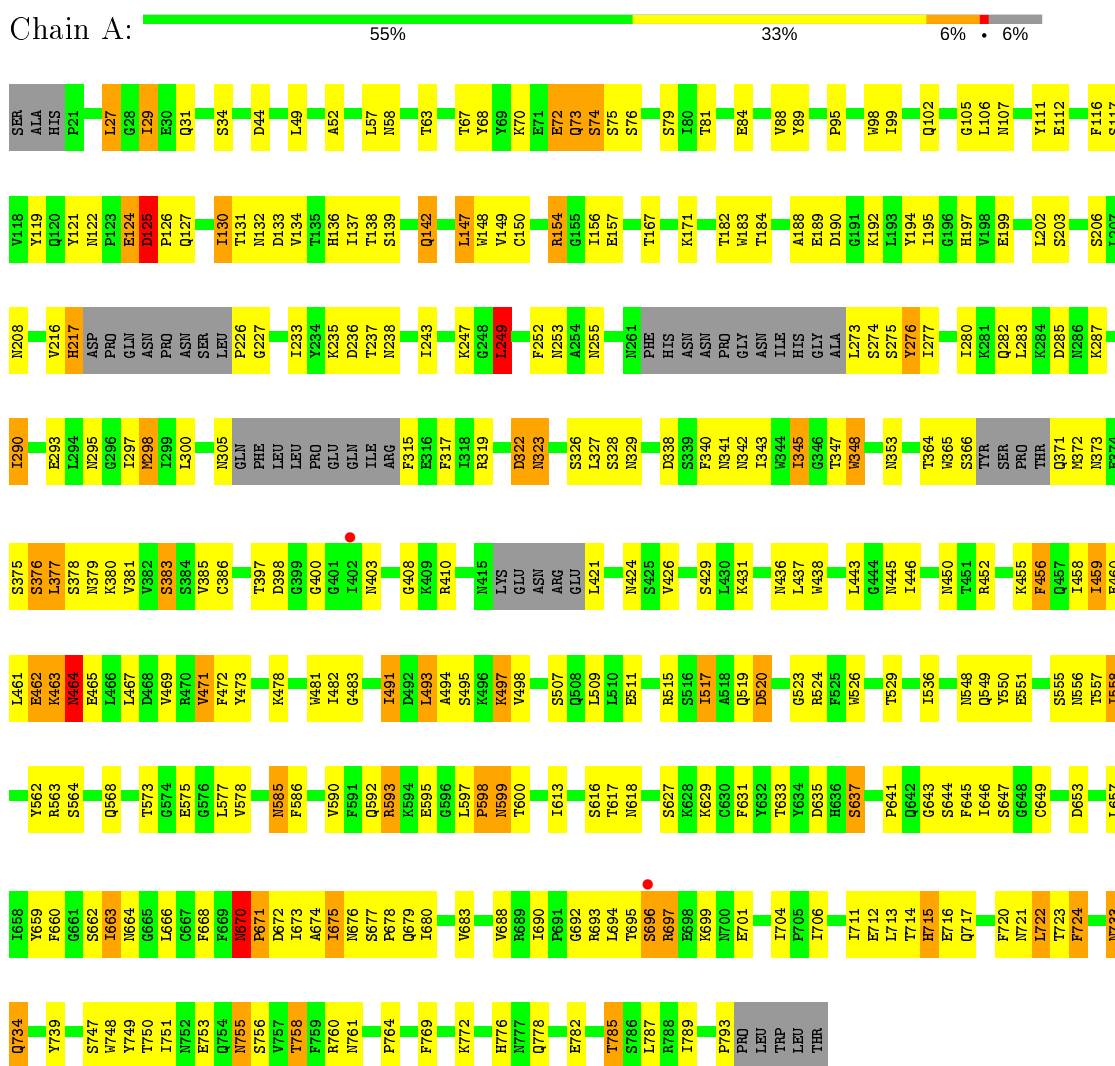
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	SER	-	EXPRESSION TAG	UNP Q8A3A5
B	18	SER	-	EXPRESSION TAG	UNP Q8A3A5
C	18	SER	-	EXPRESSION TAG	UNP Q8A3A5
D	18	SER	-	EXPRESSION TAG	UNP Q8A3A5

3 Residue-property plots

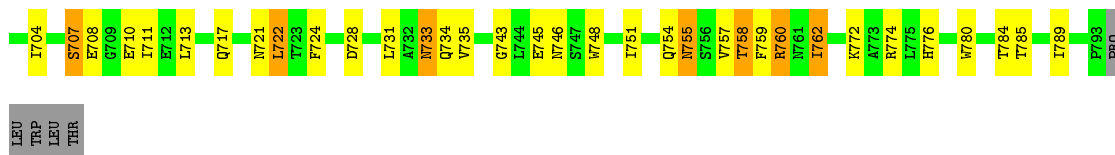
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Two-component system sensor histidine kinase/response regulator, hybrid (One-component system)



- Molecule 1: Two-component system sensor histidine kinase/response regulator, hybrid (One-component system)





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	79.93Å 114.12Å 487.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.83 – 3.30 39.83 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.6 (39.83-3.30) 99.7 (39.83-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.60 (at 3.32Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.225 , 0.277 0.228 , 0.278	Depositor DCC
R_{free} test set	3459 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	89.9	Xtrriage
Anisotropy	0.102	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 36.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	23560	wwPDB-VP
Average B, all atoms (Å ²)	82.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.06% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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

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	0.72	6/6026 (0.1%)	0.80	4/8167 (0.0%)
1	B	0.69	6/6026 (0.1%)	0.80	7/8167 (0.1%)
1	C	0.75	10/6026 (0.2%)	0.84	12/8167 (0.1%)
1	D	0.65	7/6026 (0.1%)	0.77	5/8167 (0.1%)
All	All	0.70	29/24104 (0.1%)	0.80	28/32668 (0.1%)

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	3
1	B	0	3
1	C	0	2
1	D	0	1
All	All	0	9

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	697	ARG	CZ-NH1	20.40	1.59	1.33
1	A	456	PHE	CG-CD2	16.02	1.62	1.38
1	C	169	TYR	CD1-CE1	7.43	1.50	1.39
1	A	456	PHE	CE2-CZ	7.00	1.50	1.37
1	A	348	TRP	CD2-CE2	6.55	1.49	1.41

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	169	TYR	CB-CG-CD2	12.17	128.30	121.00
1	B	249	LEU	CA-CB-CG	7.99	133.69	115.30
1	C	169	TYR	CG-CD1-CE1	7.71	127.47	121.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	456	PHE	CB-CG-CD1	-7.51	115.54	120.80
1	C	169	TYR	CD1-CG-CD2	-7.51	109.64	117.90

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	598	PRO	Peptide
1	A	672	ASP	Peptide
1	A	74	SER	Peptide
1	B	598	PRO	Peptide
1	B	670	ASN	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	5890	0	5669	267	0
1	B	5890	0	5669	242	0
1	C	5890	0	5669	246	0
1	D	5890	0	5669	236	0
All	All	23560	0	22676	964	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 21.

The worst 5 of 964 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:497:LYS:O	1:C:498:VAL:HG22	1.23	1.32
1:C:491:ILE:HD12	1:C:497:LYS:HB2	1.31	1.11
1:A:450:ASN:HD21	1:A:452:ARG:HB3	0.97	1.10
1:B:29:ILE:H	1:B:29:ILE:HD12	1.00	1.10
1:B:329:ASN:HB3	1:B:347:THR:CG2	1.84	1.07

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	724/781 (93%)	632 (87%)	76 (10%)	16 (2%)	6	30
1	B	724/781 (93%)	634 (88%)	71 (10%)	19 (3%)	5	27
1	C	724/781 (93%)	638 (88%)	69 (10%)	17 (2%)	6	29
1	D	724/781 (93%)	637 (88%)	69 (10%)	18 (2%)	5	27
All	All	2896/3124 (93%)	2541 (88%)	285 (10%)	70 (2%)	6	28

5 of 70 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	125	ASP
1	A	670	ASN
1	A	671	PRO
1	B	637	SER
1	B	670	ASN

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	649/685 (95%)	570 (88%)	79 (12%)	5	20
1	B	649/685 (95%)	585 (90%)	64 (10%)	8	28
1	C	649/685 (95%)	579 (89%)	70 (11%)	6	24
1	D	649/685 (95%)	579 (89%)	70 (11%)	6	24
All	All	2596/2740 (95%)	2313 (89%)	283 (11%)	6	24

5 of 283 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	629	LYS
1	C	285	ASP
1	D	610	ASN
1	B	680	ILE
1	C	29	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 123 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	754	GLN
1	C	341	ASN
1	D	638	ASN
1	C	35	ASN
1	C	240	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](#)

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

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	731/781 (93%)	-0.07	2 (0%) 94 94	39, 71, 121, 177	0
1	B	731/781 (93%)	0.01	9 (1%) 79 78	44, 75, 135, 210	0
1	C	731/781 (93%)	0.04	14 (1%) 66 65	38, 80, 147, 221	0
1	D	731/781 (93%)	0.02	12 (1%) 72 70	43, 77, 149, 219	0
All	All	2924/3124 (93%)	-0.00	37 (1%) 77 77	38, 76, 139, 221	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	699	LYS	5.7
1	D	700	ASN	5.4
1	D	697	ARG	4.9
1	C	697	ARG	4.3
1	D	698	GLU	4.2

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.