Full wwPDB X-ray Structure Validation Report (i)

Jun 1, 2026 – 08:03 AM EDT

PDB ID	:	9OZ4 / pdb_00009oz4
Title	:	Crystal Structure of Substrate Binding Protein (TAXI-TRAP) in Complex with L-Glutamate from Bordetella pertussis
Authors	:	Antony, I.R.; Dobson, R.
Deposited on	:	2025-06-05
Resolution	:	2.01 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (i)) were used in the production of this report:

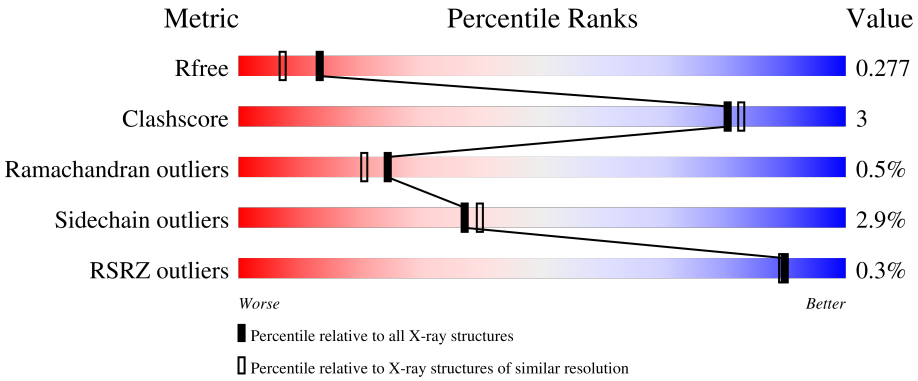
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.01 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	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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 <=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	314	88% 6% • 5%
1	B	314	% 88% 6% • 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9186 atoms, of which 4462 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 TRAP transporter solute receptor, TAXI family.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	299	Total	C	H	N	O	S	61	0	0
			4482	1439	2225	381	429	8			
1	B	299	Total	C	H	N	O	S	61	0	0
			4482	1439	2225	381	429	8			

There are 32 discrepancies between the modelled and reference sequences:

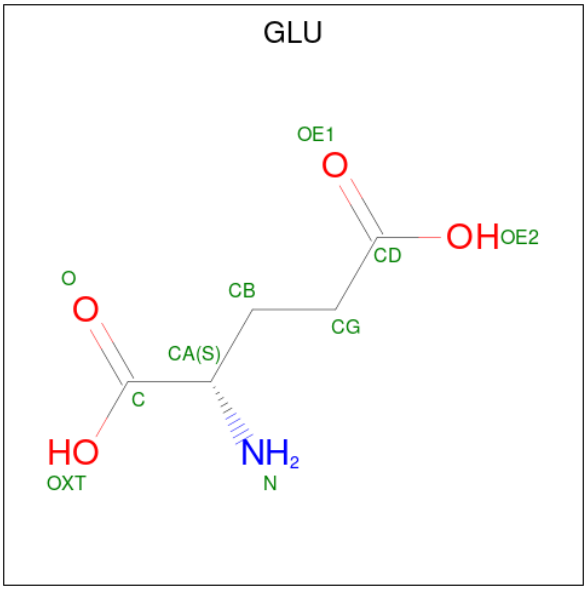
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP A0A381A3N1
A	300	GLU	-	expression tag	UNP A0A381A3N1
A	301	ASN	-	expression tag	UNP A0A381A3N1
A	302	LEU	-	expression tag	UNP A0A381A3N1
A	303	TYR	-	expression tag	UNP A0A381A3N1
A	304	PHE	-	expression tag	UNP A0A381A3N1
A	305	GLN	-	expression tag	UNP A0A381A3N1
A	306	SER	-	expression tag	UNP A0A381A3N1
A	307	ALA	-	expression tag	UNP A0A381A3N1
A	308	GLY	-	expression tag	UNP A0A381A3N1
A	309	HIS	-	expression tag	UNP A0A381A3N1
A	310	HIS	-	expression tag	UNP A0A381A3N1
A	311	HIS	-	expression tag	UNP A0A381A3N1
A	312	HIS	-	expression tag	UNP A0A381A3N1
A	313	HIS	-	expression tag	UNP A0A381A3N1
A	314	HIS	-	expression tag	UNP A0A381A3N1
B	1	MET	-	initiating methionine	UNP A0A381A3N1
B	300	GLU	-	expression tag	UNP A0A381A3N1
B	301	ASN	-	expression tag	UNP A0A381A3N1
B	302	LEU	-	expression tag	UNP A0A381A3N1
B	303	TYR	-	expression tag	UNP A0A381A3N1
B	304	PHE	-	expression tag	UNP A0A381A3N1
B	305	GLN	-	expression tag	UNP A0A381A3N1
B	306	SER	-	expression tag	UNP A0A381A3N1
B	307	ALA	-	expression tag	UNP A0A381A3N1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	308	GLY	-	expression tag	UNP A0A381A3N1
B	309	HIS	-	expression tag	UNP A0A381A3N1
B	310	HIS	-	expression tag	UNP A0A381A3N1
B	311	HIS	-	expression tag	UNP A0A381A3N1
B	312	HIS	-	expression tag	UNP A0A381A3N1
B	313	HIS	-	expression tag	UNP A0A381A3N1
B	314	HIS	-	expression tag	UNP A0A381A3N1

- Molecule 2 is GLUTAMIC ACID (CCD ID: GLU) (formula: $C_5H_9NO_4$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	0	0
			16	5	6	1	4		
2	B	1	Total	C	H	N	O	0	0
			16	5	6	1	4		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	132	Total	O	0	0
			132	132		
3	B	58	Total	O	0	0
			58	58		

3 Residue-property plots (i)

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: TRAP transporter solute receptor, TAXI family

Chain A:

• Molecule 1: TRAP transporter solute receptor, TAXI family

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.43Å 76.91Å 65.23Å 90.00° 93.68° 90.00°	Depositor
Resolution (Å)	46.00 – 2.01 46.00 – 2.01	Depositor EDS
% Data completeness (in resolution range)	99.4 (46.00-2.01) 98.7 (46.00-2.01)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.8.0267, REFMAC 5.8.0267	Depositor
R, R_{free}	0.165 , 0.278 0.164 , 0.277	Depositor DCC
R_{free} test set	1805 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å^2)	30.1	Xtriage
Anisotropy	0.781	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 48.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	9186	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.32% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	1.08	1/2310 (0.0%)	1.37	1/3127 (0.0%)
1	B	1.14	2/2310 (0.1%)	1.43	1/3127 (0.0%)
All	All	1.11	3/4620 (0.1%)	1.40	2/6254 (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	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	295	GLU	CD-OE1	12.13	1.48	1.25
1	B	295	GLU	CD-OE2	9.46	1.43	1.25
1	A	132	SER	C-O	6.00	1.31	1.24

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	83	GLY	CA-C-O	-6.06	118.28	122.22
1	B	91	VAL	N-CA-C	-5.67	106.15	111.48

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	295	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	2257	2225	2217	13	0
1	B	2257	2225	2217	10	0
2	A	10	6	5	1	0
2	B	10	6	5	1	0
3	A	132	0	0	2	0
3	B	58	0	0	0	0
All	All	4724	4462	4444	23	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 3.

All (23) 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:158:HIS:HE1	1:A:180:ASN:HD21	1.17	0.91
1:B:158:HIS:HE1	1:B:180:ASN:HD21	1.21	0.84
1:A:158:HIS:CE1	1:A:180:ASN:HD21	2.07	0.70
1:A:59:LEU:HD13	3:A:600:HOH:O	1.94	0.67
1:A:65:GLN:HE21	2:A:401:GLU:N	1.95	0.64
1:A:158:HIS:HE1	1:A:180:ASN:ND2	1.97	0.58
1:B:158:HIS:CE1	1:B:180:ASN:HD21	2.12	0.58
1:B:75:GLY:O	1:B:80:LYS:HA	2.04	0.57
1:B:102:LEU:CD1	1:B:192:LEU:HD23	2.35	0.57
1:B:158:HIS:HE1	1:B:180:ASN:ND2	1.96	0.56
1:A:133:MET:HE3	1:A:133:MET:HA	1.89	0.54
1:B:65:GLN:HE21	2:B:401:GLU:N	2.05	0.53
1:A:155:PRO:HA	1:A:158:HIS:CD2	2.47	0.49
1:B:260:ASP:O	1:B:263:LYS:HB2	2.15	0.46
1:A:104:ARG:HD3	1:A:168:ASP:HA	1.97	0.46
1:A:6:VAL:HG13	1:A:61:LEU:HD12	1.97	0.46
1:A:6:VAL:HG13	1:A:61:LEU:CD1	2.47	0.44
1:A:186:THR:OG1	1:A:221:PRO:HG2	2.18	0.43
1:A:61:LEU:HD23	1:A:238:THR:HB	1.99	0.43
1:A:4:LYS:HE3	3:A:531:HOH:O	2.18	0.42
1:B:298:LEU:C	1:B:299:LEU:HG	2.45	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:6:VAL:HG12	1:B:39:CYS:HA	2.02	0.42
1:B:292:TYR:CD1	1:B:292:TYR:C	2.99	0.40

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	297/314 (95%)	289 (97%)	7 (2%)	1 (0%)	36 35
1	B	297/314 (95%)	286 (96%)	9 (3%)	2 (1%)	18 14
All	All	594/628 (95%)	575 (97%)	16 (3%)	3 (0%)	24 21

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	3	GLN
1	A	94	ILE
1	B	94	ILE

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	237/251 (94%)	229 (97%)	8 (3%)	32 33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	237/251 (94%)	231 (98%)	6 (2%)	42 45
All	All	474/502 (94%)	460 (97%)	14 (3%)	37 38

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	GLN
1	A	6	VAL
1	A	61	LEU
1	A	110	LYS
1	A	133	MET
1	A	145	LYS
1	A	195	LEU
1	A	259	GLU
1	B	6	VAL
1	B	26	ARG
1	B	42	GLU
1	B	61	LEU
1	B	102	LEU
1	B	195	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	158	HIS
1	A	245	GLN
1	B	158	HIS
1	B	180	ASN
1	B	182	GLN
1	B	290	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 oligosaccharides 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	$\#\|Z\| > 2$
2	GLU	B	401	-	8,9,9	1.06	0	8,11,11	1.47	2 (25%)
2	GLU	A	401	-	8,9,9	1.22	2 (25%)	8,11,11	0.96	0

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
2	GLU	B	401	-	-	2/9/9/9	-
2	GLU	A	401	-	-	2/9/9/9	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	GLU	OXT-C	-2.41	1.23	1.30
2	A	401	GLU	OE2-CD	-2.05	1.24	1.30

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	GLU	OXT-C-O	-2.72	117.91	124.08
2	B	401	GLU	OE1-CD-CG	-2.02	116.69	123.09

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	401	GLU	O-C-CA-N
2	A	401	GLU	OXT-C-CA-N
2	A	401	GLU	O-C-CA-N
2	B	401	GLU	OXT-C-CA-N

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	GLU	1	0
2	A	401	GLU	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

6.1 Protein, DNA and RNA chains

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^{th} 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	299/314 (95%)	-0.69	0	100	100	21, 33, 52, 63	0
1	B	299/314 (95%)	-0.49	2 (0%)	84	84	26, 43, 69, 122	0
All	All	598/628 (95%)	-0.59	2 (0%)	90	89	21, 37, 64, 122	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	299	LEU	3.4
1	B	1	MET	2.5

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
2	GLU	A	401	10/10	0.98	0.04	21,25,33,35	0
2	GLU	B	401	10/10	0.98	0.04	27,35,41,43	0

6.5 Other polymers (i)

There are no such residues in this entry.