



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 04:50 AM UTC

PDB ID : 8BT6 / pdb_00008bt6
Title : Crystal structure of Toxoplasma gondii glideosome-associated connector
Authors : Hung, Y.F.; Kursula, I.
Deposited on : 2022-11-27
Resolution : 2.33 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
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

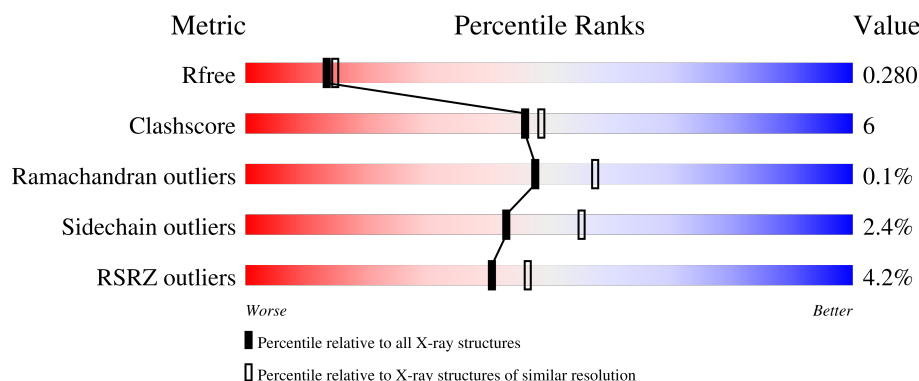
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.33 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	3031 (2.36-2.32)
Clashscore	190562	3127 (2.36-2.32)
Ramachandran outliers	187476	3095 (2.36-2.32)
Sidechain outliers	187428	3095 (2.36-2.32)
RSRZ outliers	180081	3033 (2.36-2.32)

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	2646	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ACT	A	2701	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 40007 atoms, of which 19985 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.

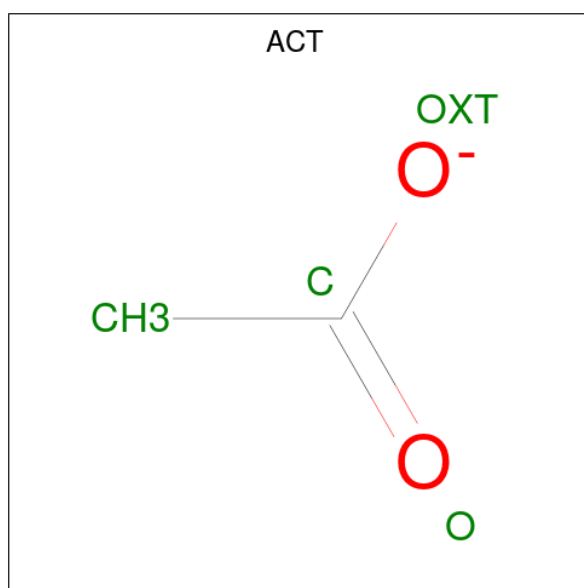
- Molecule 1 is a protein called Putative anonymous antigen-1.

Mol	Chain	Residues	Atoms								ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S	Se				
1	A	2633	39936	12477	19982	3407	3905	75	90		0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP A0A151HKA5
A	2640	HIS	-	expression tag	UNP A0A151HKA5
A	2641	HIS	-	expression tag	UNP A0A151HKA5
A	2642	HIS	-	expression tag	UNP A0A151HKA5
A	2643	HIS	-	expression tag	UNP A0A151HKA5
A	2644	HIS	-	expression tag	UNP A0A151HKA5
A	2645	HIS	-	expression tag	UNP A0A151HKA5

- Molecule 2 is ACETATE ION (CCD ID: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 3 is POTASSIUM ION (CCD ID: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	K	0	0
			2	2		

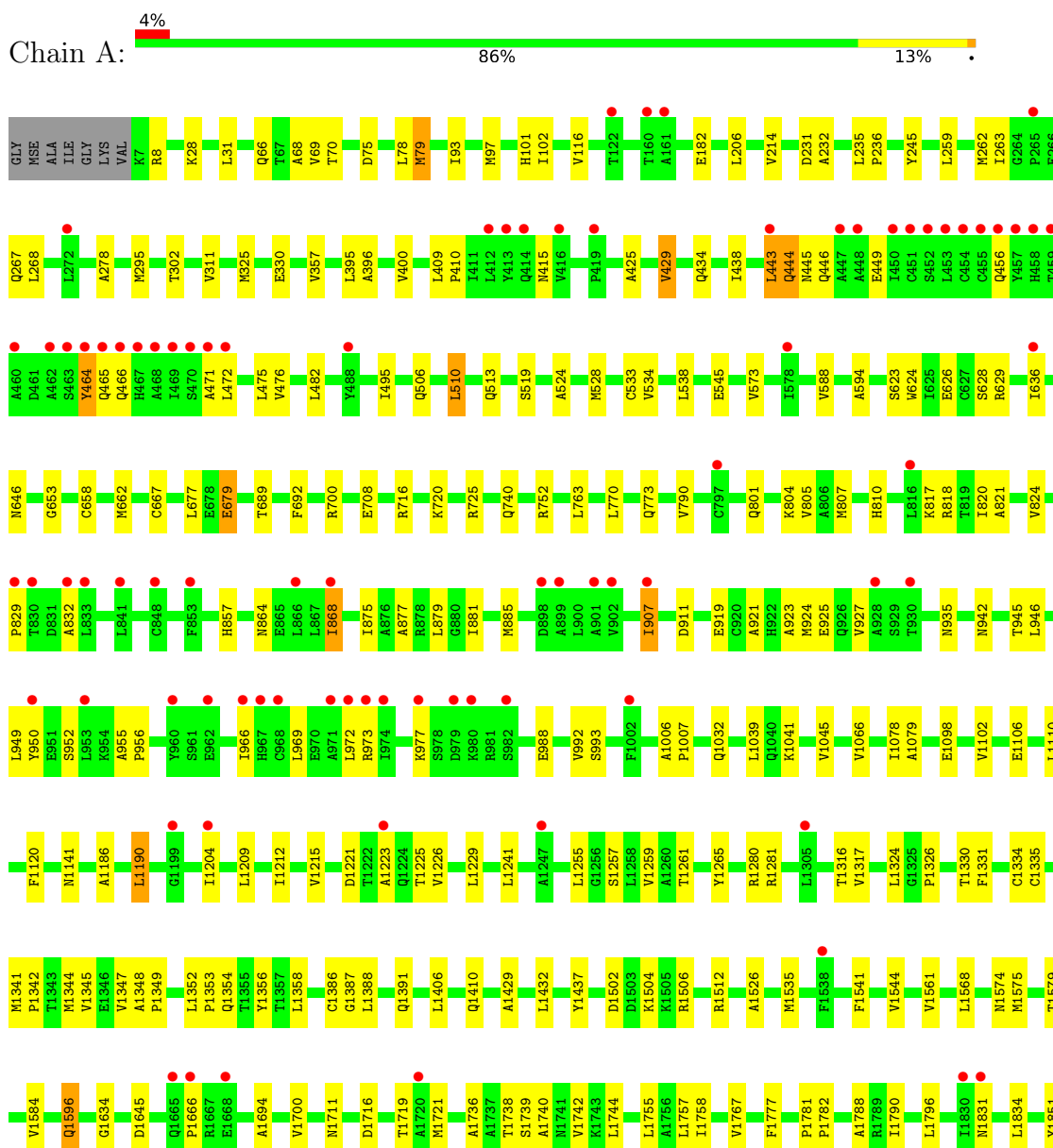
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	62	Total	O	0	0
			62	62		

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 anonymous antigen-1



V2629	D2432	I2177	T2016	D1852
Q2630	F2481	I2190	N2017	K1853
V2631	F2484	M2193	G2018	G1854
A2635	T2487	M2194	A2019	L1861
F2639	D2492	V2215	P2020	R1862
HIS	K2520	E2238	L2023	V1872
HIS	V2526	I2245	T2037	V1877
HIS	F2527	A2249	A2038	L1880
HIS	L2528	I2252	L2039	L1886
HIS	P2529	S2255	E2043	Q1887
HIS	W2538	P2259	S2050	R1888
	W2549	M2262	T2055	L1889
	C2550	L2263	T2056	I1890
	M2551	Q2284	P2057	V1896
	G2552		N2058	M1900
	N2553		L2059	L1903
	T2563		Y2071	L1903
	V2564	A2311	L2078	L1913
	R2565	G2312	L2087	I1919
	I2566	V2313	T2093	K1922
	R2567	V2314	L2094	L1926
	N2568	V2328	V2095	L1927
	V2569	E2331	M2099	G1928
	V2574	A2335	E2100	L1935
	K2584		N2101	L1938
	V2589	M2345	E2102	K1942
	T2594	D2346	M2106	K1942
	M2595	E2347	C2116	V1948
	C2596	V2348	M2117	
	L2597	M2349	C2120	A1957
	N2602	I2350	L2120	C1960
	D2603	L2354	D2126	I1961
	D2604	H2366	T2127	G1962
	F2605	F2383	L2134	V1982
	L2611	G2389	D2135	L1983
	P2612	M2406	I2136	N1984
	M2613	M2416	R2142	D1985
	V2614	M2419	N2153	N1994
	F2624	S2419	E2157	V1998
	V2625	I2423	L2172	L1999
	E2626		L2173	N2002
	M2627			
	M2628			

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	121.40Å 146.66Å 185.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.33 – 2.33 47.33 – 2.33	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.33-2.33) 99.9 (47.33-2.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.20_4459	Depositor
R, R_{free}	0.229 , 0.272 0.243 , 0.280	Depositor DCC
R_{free} test set	7088 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	55.7	Xtriage
Anisotropy	0.492	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 50.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	40007	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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](#)

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

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	0.11	0/20146	0.29	0/27155

There are no bond length outliers.

There are no bond angle outliers.

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	19954	19982	19982	232	1
2	A	4	3	3	0	0
3	A	2	0	0	0	0
4	A	62	0	0	14	0
All	All	20022	19985	19985	232	1

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 6.

The worst 5 of 232 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:2345:MSE:HE3	1:A:2349:MSE:HE2	1.32	1.11
1:A:519:SER:HA	1:A:528:MSE:HE1	1.52	0.89
1:A:524:ALA:HB1	1:A:528:MSE:HE2	1.55	0.86
1:A:513:GLN:OE1	4:A:2801:HOH:O	1.94	0.84
1:A:817:LYS:NZ	4:A:2805:HOH:O	2.12	0.80

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1437:TYR:OH	1:A:2520:LYS:NZ[3_655]	2.18	0.02

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	2631/2646 (99%)	2542 (97%)	87 (3%)	2 (0%)	48 57

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	415	ASN
1	A	1666	PRO

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	2145/2064 (104%)	2093 (98%)	52 (2%)	43 55

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	907	ILE
1	A	1410	GLN
1	A	2549	TRP
1	A	911	ASP
1	A	1190	LEU

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

Mol	Chain	Res	Type
1	A	2101	ASN
1	A	2400	GLN
1	A	2559	ASN
1	A	2408	ASN
1	A	2257	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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	ACT	A	2701	-	3,3,3	1.17	0	3,3,3	1.47	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	2543/2646 (96%)	0.37	107 (4%)	40 46	50, 81, 128, 182	0

The worst 5 of 107 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	454	CYS	6.8
1	A	448	ALA	6.3
1	A	1199	GLY	6.1
1	A	462	ALA	6.1
1	A	2639	PHE	5.6

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](#)

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, 95th 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	ACT	A	2701	4/4	0.72	0.58	95,114,137,137	0
3	K	A	2702	1/1	0.94	0.16	81,81,81,81	0
3	K	A	2703	1/1	0.94	0.06	81,81,81,81	0

6.5 Other polymers [i](#)

There are no such residues in this entry.