



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 05:34 AM UTC

PDB ID : 8C4A / pdb_00008c4a
Title : Structural and interactional insights into the glideosome-associated connector from *Toxoplasma gondii*
Authors : Kumar, A.; Morgan, R.M.L.; Matthews, S.J.
Deposited on : 2023-01-03
Resolution : 2.67 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

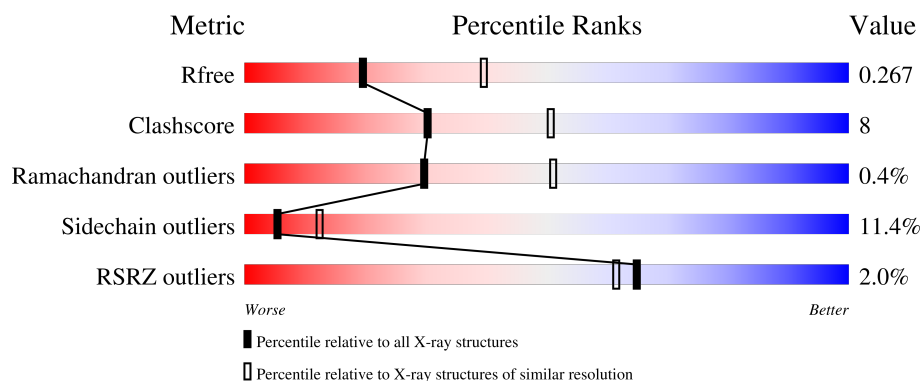
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.67 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	5070 (2.70-2.66)
Clashscore	190562	5409 (2.70-2.66)
Ramachandran outliers	187476	5324 (2.70-2.66)
Sidechain outliers	187428	5324 (2.70-2.66)
RSRZ outliers	180081	5070 (2.70-2.66)

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	2498	2% 75% 19% ..

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 37731 atoms, of which 18940 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
1	A	2474	Total	C	H	N	O	S	Se		0	2	0
			37395	11670	18716	3181	3671	73	84				

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	564	GLU	ALA	conflict	UNP A0A7J6JYP1

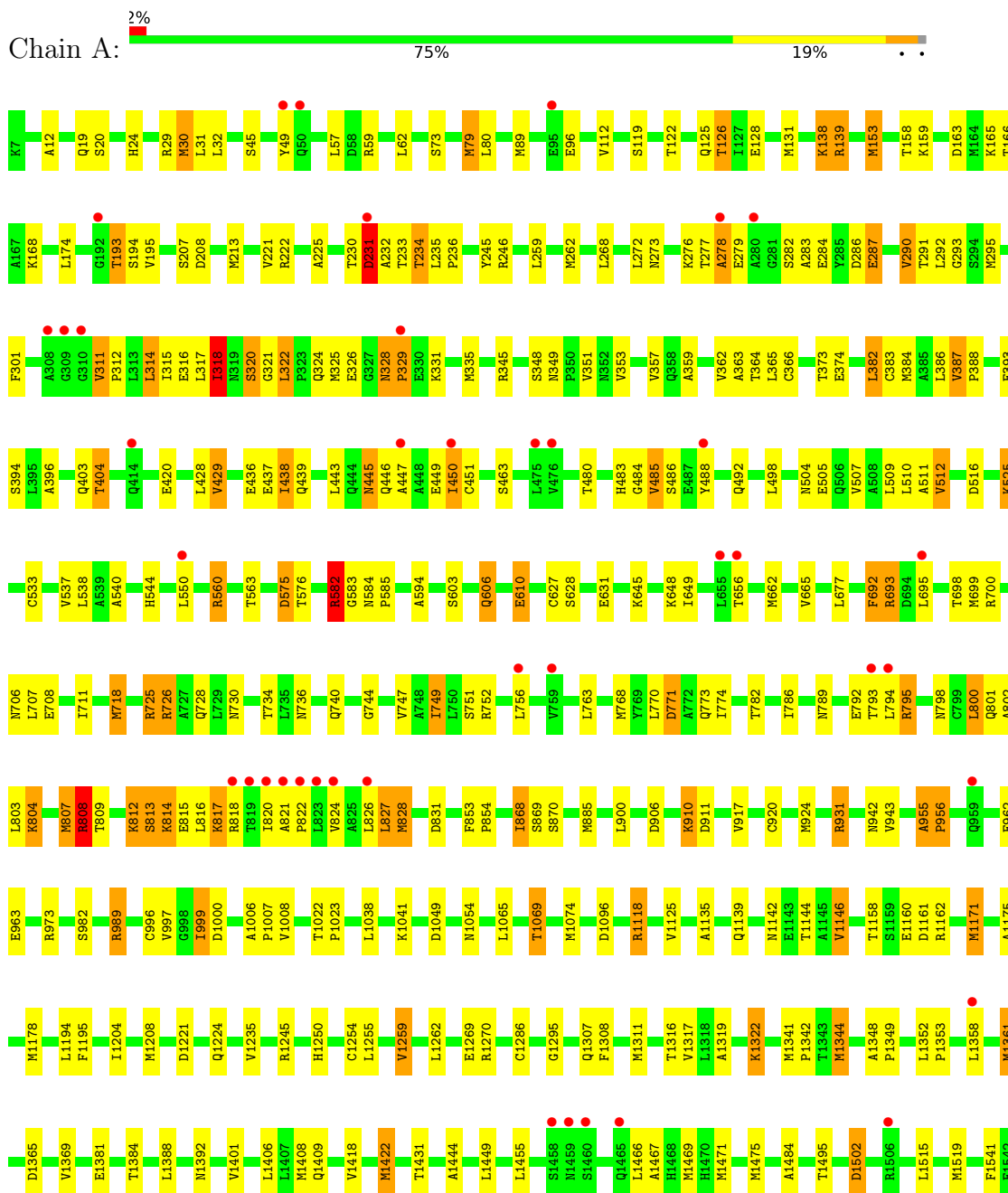
- Molecule 2 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	112	Total	H	O	0	0
			336	224	112		

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: Putative anonymous antigen-1



K2464	PRO	R2227	K2036	L1889	H1725	V1543
C2467	PRO	R2228	T2037	I1893	K1743	M1554
T2470	GLU	V2229	A2038	I1893	L1744	I1558
D2479	GLU	E2234	R2041	E1918	S1751	Q1566
T2489	GLU	L2235	L2042	I1919	E1754	G1567
N2499	GLU	A2236	E2043	K1920	I1758	L1568
V2500	GLU	C2240	S2044	T1921	E1765	S1569
P2504	GLU	G2155	V2045	L1936	I1766	M1570
		I2156	A2048	L1940	E1785	A1571
		E2167	I2049	P1941	I1787	D1572
		T2247	S2050	A1946	V1767	M1575
		V2251	S2053	T1950	F1780	R1578
		L2256	L2054	T1951	P1781	E1582
		P2259	T2056	V1952	A1784	T1592
		M2260	I2059	I1961	Q1785	T1593
		V2261	Q2060	G1962	D1786	R1620
		M2262	P2061	H1963	T1787	L1623
		L2263	F2062	K1964	A1788	G1627
		E2264	L2063	K1965	R1789	T1631
		A2265	I2067	I1969	I1790	E1636
		C2266	Q2086	L1978	Q1791	E1652
		R2267	L2087	L1978	G1792	K1653
		V2268	E2088	V1982	C1793	V1654
		L2272	R2092	V1992	P1794	A1655
		L2273	T2093	C1993	R1813	M1656
		R2278	L2094	C2001	H1817	M1661
		V2287	M2099	L2004	R1821	E1662
		L2289	N2105	L2004	Y1822	R1681
		Q2295	M2106	L2004	V1825	T1690
		R2297	R2107	T2005	N1831	L1703
		A2298	K2108	N2007	R1832	M1707
		M2301	V2111	E2008	Y1835	G1708
		HIS	C2112	S2009	K1838	M1709
		ALA	L2113	M2010	T1838	
		GLN	L2114	K2011	P1839	
		ALA	P2115	L2012	I1840	
		GLN	C2116	L2013	M1841	
		ALA	M2117	L2014	G1708	
		ALA	D2217	N2017	K1853	
		ASP	V2118	G2018	L1715	
		TYR	V2219	A2019	D1716	
		GLU	V2220	P2020	L1861	
		ALA	G2221	E1879	E1879	
		GLY	R2124	L1880	L1719	
		VAL	L2223	S2028	A1720	
		VAL	R2224	D2031	M1721	
		SER	C2225	L2130	V1885	
		GLU	T2226	E2035	G1722	

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	119.08Å 123.61Å 221.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	110.75 – 2.67 110.75 – 2.67	Depositor EDS
% Data completeness (in resolution range)	100.0 (110.75-2.67) 99.9 (110.75-2.67)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.53 (at 2.69Å)	Xtriage
Refinement program	REFMAC v5.0	Depositor
R, R_{free}	0.209 , 0.268 0.232 , 0.267	Depositor DCC
R_{free} test set	4517 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å ²)	85.0	Xtriage
Anisotropy	0.013	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 35.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	37731	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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

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	0.56	2/18847 (0.0%)	1.02	10/25412 (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	A	0	24

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	492[A]	GLN	C-O	5.72	1.31	1.24
1	A	492[B]	GLN	C-O	5.72	1.31	1.24

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2479	ASP	CA-CB-CG	7.76	120.36	112.60
1	A	692	PHE	CA-CB-CG	6.59	120.39	113.80
1	A	1144	THR	CA-CB-OG1	-5.71	101.04	109.60
1	A	585	PRO	N-CA-C	-5.57	106.90	113.86
1	A	516	ASP	CA-CB-CG	5.27	117.87	112.60

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	139	ARG	Sidechain
1	A	345	ARG	Sidechain
1	A	560	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	582	ARG	Sidechain
1	A	59	ARG	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	18679	18716	18693	314	0
2	A	112	224	0	3	0
All	All	18791	18940	18693	314	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 8.

The worst 5 of 314 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:807:MSE:SE	1:A:868:ILE:HD13	1.85	1.27
1:A:1307:GLN:HG3	1:A:1311:MSE:HE2	1.18	1.07
1:A:885:MSE:HE1	1:A:920:CYS:SG	2.13	0.89
1:A:955:ALA:HB1	1:A:956:PRO:HD2	1.55	0.87
1:A:885:MSE:HE2	1:A:924:MSE:HE2	1.58	0.85

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	2470/2498 (99%)	2312 (94%)	147 (6%)	11 (0%)	30 51

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	827	LEU
1	A	318	ILE
1	A	348	SER
1	A	956	PRO
1	A	329	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	2007/1944 (103%)	1778 (89%)	229 (11%)	5 12

5 of 229 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	910	LYS
1	A	2404	GLN
1	A	1593	THR
1	A	2371	ARG
1	A	2160	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2096	ASN
1	A	2104	ASN
1	A	2369	GLN
1	A	1076	ASN
1	A	959	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](#)

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	2390/2498 (95%)	0.02	48 (2%) 65 61	36, 74, 97, 121	2 (0%)

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	822	PRO	5.6
1	A	476	VAL	4.0
1	A	475	LEU	3.7
1	A	329	PRO	3.7
1	A	95	GLU	3.7

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.