



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 26, 2024 – 01:14 AM EDT

PDB ID : 6SUQ
Title : Crystal Structure of TcdB2-TccC3-TEV
Authors : Roderer, D.; Schubert, E.; Sitsel, O.; Raunser, S.
Deposited on : 2019-09-16
Resolution : 3.70 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	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

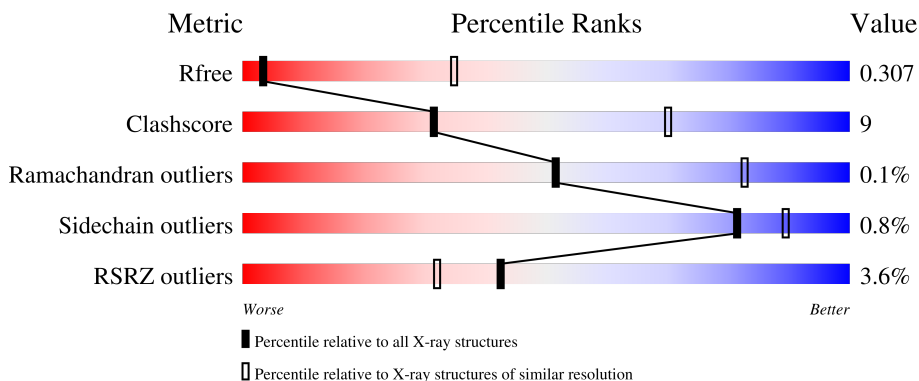
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)
RSRZ outliers	127900	1578 (3.90-3.50)

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	2410	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 17025 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 TcdB2,TccC3,Genome polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	2134	17025	10667	3024	3300	34	0	0	0

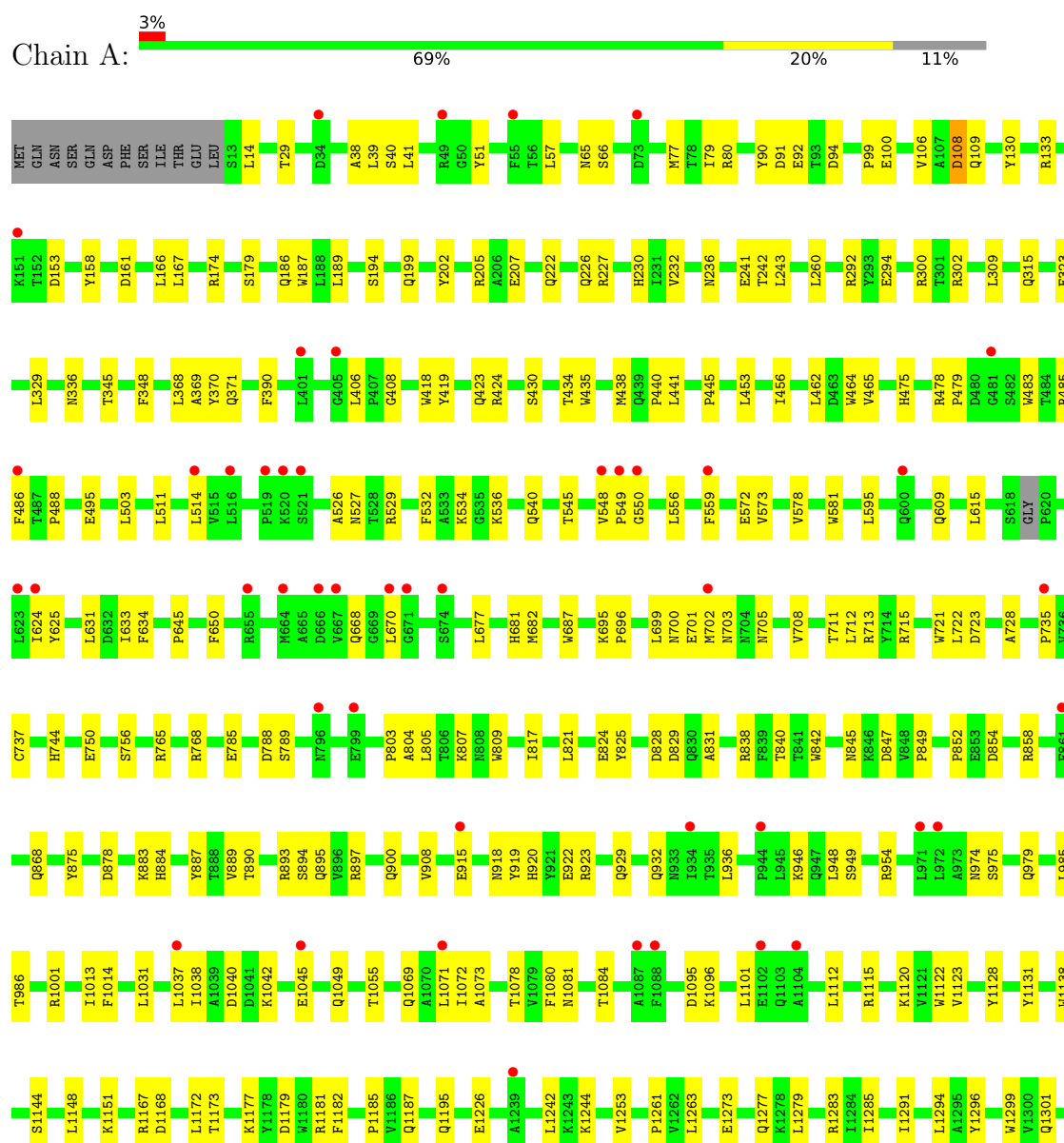
There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1475	PRO	-	linker	UNP Q8GF99
A	1476	GLY	-	linker	UNP Q8GF99
A	1477	SER	-	linker	UNP Q8GF99
A	1478	ARG	-	linker	UNP Q8GF99
A	1479	PRO	-	linker	UNP Q8GF99
A	2160	GLU	-	linker	UNP Q8GF97
A	2161	PHE	-	linker	UNP Q8GF97
A	2380	VAL	SER	conflict	UNP P04517
A	2398	ARG	-	expression tag	UNP P04517
A	2399	ARG	-	expression tag	UNP P04517
A	2400	ARG	-	expression tag	UNP P04517
A	2401	ARG	-	expression tag	UNP P04517
A	2402	ARG	-	expression tag	UNP P04517
A	2403	LEU	-	expression tag	UNP P04517
A	2404	GLU	-	expression tag	UNP P04517
A	2405	HIS	-	expression tag	UNP P04517
A	2406	HIS	-	expression tag	UNP P04517
A	2407	HIS	-	expression tag	UNP P04517
A	2408	HIS	-	expression tag	UNP P04517
A	2409	HIS	-	expression tag	UNP P04517
A	2410	HIS	-	expression tag	UNP P04517

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: TcdB2,TccC3,Genome polypeptide



HIS	HIS	THR	ASN	PRO	S9081	L1958	P1806	L1662	R1505	N1408	K1304
LEU	LEU	SER	THR	ARG	Q2062	V1959	M1809	L1668	R1506	D1409	P1307
THR	THR	THR	THR	TYR	G2069	K1960	V1810	V1668	N1507	W1410	
ASN	ASN	ASN	LEU	ASN	Q2070	D1962	I1811	S1673	D1508	R1411	I1310
GLN	GLN	PRO	GLN	PRO	I2071	K1963	S1812	Q1674	F1509	Y1412	
GLU	SER	SER	GLN	ILE	E2076	D1967	I1813	L1675	H1510	M1415	L1313
ALA	SER	SER	HIS	SER	Y2077	D1968	R1814	L1676	N1314	N1315	
GLN	ASP	GLY	LEU	THR	Y2078	D1969	W1823	Y1677	D1519	R1419	
GLN	THR	ILE	ASP	ILE	P2079	R1969	Q1826	D1678	T1520	K1422	
VAL	PHE	GLY	GLY	CYS	Y2080	R1973	Q1826	T1679	R1521		L1319
SER	TRP	ARG	THR	HIS	Y1974	S1975	Y1836	Q1680	R1524	L1324	
GLY	LYS	ASP	LEU	LEU	L2085	S1976		L1696	H1525		
TRP	HIS	MET	THR	THR	W2086	D1977	L1842	A1697	Q1526	D1329	
ARG	TRP	ILE	ASN	ASN	A2087	G1976		T1428	Y1527	R1330	
LEU	LEU	ILE	ILE	GLU	A2088	D1977	R1848	Y1700	D1528	Y1331	
ASN	GLN	ILE	ILE	SER	R2089	M1981	E1849	Y1701	H1529		P1335
ALA	THR	ARG	ASP	ASP	N2090	N1992	F1860	T1715	H1530	D1432	
ASP	LYS	MET	GLY	GLY	E2093	T1999	A1864	Q1724	Y1431	P1433	Q1337
SER	ASP	PRO	HIS	HIS	T2098	T1999		N1722	D1436		G1348
VAL	GLY	THR	LYS	THR	I2099	L2004		Q1725	R1437		
LEU	GLN	THR	ASP	THR	E2005	L2006	D1882	R1726	E1438		R1351
TRP	CYS	PHE	SER	PHE	R2100	R2007	R1883	A1735	N1564		
GLY	GLY	PRO	PRO	TYR	Y2101			E1746	D1572		H1359
SER	SER	PHE	PHE	GLY	E2105	T2015	H1892	Q1746	R1575		G1362
HIS	LYS	LEU	GLN	ILE	T2109	L2019	I1905	N1735	I1582		M1363
VAL	LYS	THR	LYS	PHE	Y2112	Q2020	H1912	E1746	V1590		A1364
MET	THR	ARG	LEU	PRO	Y2112	Q2021	R1913	L1753	T1591		R1365
VAL	ARG	ASP	PHE	PHE	W2122	V2021	A1914	A1757	A1592		Q1366
PRO	GLY	GLY	ARG	ILE	I2123	T2023	V1915	K1761	I1596		L1373
GLU	PHE	GLU	GLU	ILE	G2124	E2026	L1919	L1762	Q1597		I1374
PRO	VAL	ILE	ASN	THR	R2125	G2027	D1922	R1763	VAL		T1380
VAL	GLY	GLY	ARG	LYS	T2134	G2028	Q1925	E1764	LYS		E1381
PHE	GLY	ILE	GLU	HIS	L2140	R2033	Q1925	H1766	MET		N1382
GLN	ILE	HIS	GLU	LEU	V2144	R2034	D1931	I1771	PRO		R1383
PRO	VAL	PRO	ARG	LEU	P2148	W2037	A1932		SER		W1384
PHE	ALA	ASN	CYS	ASN	P2148	E2038	G1933		ARG		A1385
THR	PHE	VAL	LEU	ASN	P2154	K2041	G1934		PRO		
GLN	THR	THR	THR	THR	L2157	P2042	H1935		MET		T1390
LEU	THR	THR	ASN	LEU	MET	E2043	L1939		LYS		E1391
LEU	THR	THR	LEU	VAL	PRO	D2044	L1945		R1482		Y1392
ASN	ASN	ASN	GLN	VAL	THR	I2045			K1486		D1393
ARG	THR	THR	THR	GLN	GLU	N2048	N1948				Q1397
ARG	PHE	LYS	SER	SER	GLU	Q2049			Q1489		P1398
ARG	THR	THR	LEU	LEU	PHE	R2050			I1399		I1399
VAL	THR	THR	GLY	HIS	GLY	L2051	E1952		R1490		R1400
LEU	VAL	VAL	GLY	GLY	GLU	R2051	L1953		T1491		T1401
GLU	PRO	SER	SER	VAL	SER	M2056	I1954		V1494		Y1402
HIS	LYS	MET	VAL	PHE	LEU	N2056	R1801		S1495		
HIS	ASN	VAL	PHE	LYS	PHE	Y1956	Y1802		V1496		Y1405
HIS	SER	VAL	VAL	LYS	LYS	T1957	E1803				F1406
HIS	MET	ASP	LYS	LYS	GLY		Y1805		R1658		L1407

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	234.43Å 234.43Å 143.19Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.73 – 3.70 49.18 – 3.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.73-3.70) 100.0 (49.18-3.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.01 (at 3.67Å)	Xtriage
Refinement program	PHENIX 1.16-3549	Depositor
R, R_{free}	0.269 , 0.302 0.286 , 0.307	Depositor DCC
R_{free} test set	2431 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	130.9	Xtriage
Anisotropy	0.362	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 96.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.196 for -h,-k,l	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	17025	wwPDB-VP
Average B, all atoms (Å ²)	135.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% 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	0.26	0/17444	0.46	2/23779 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1407	LEU	CA-CB-CG	6.46	130.16	115.30
1	A	1407	LEU	CB-CG-CD2	-6.00	100.80	111.00

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	17025	0	16395	302	0
All	All	17025	0	16395	302	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 9.

The worst 5 of 302 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:1330:ARG:HD3	1:A:1335:PRO:HB3	1.62	0.81

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:371:GLN:HB3	1:A:696:PRO:HA	1.66	0.76
1:A:534:LYS:HD3	1:A:536:LYS:HD3	1.71	0.72
1:A:668:GLN:HB3	1:A:670:LEU:HD13	1.73	0.70
1:A:91:ASP:OD1	1:A:92:GLU:N	2.23	0.70

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	2128/2410 (88%)	2033 (96%)	92 (4%)	3 (0%)	51	83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1335	PRO
1	A	2154	PRO
1	A	479	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	1842/2095 (88%)	1828 (99%)	14 (1%)	81	89

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1382	ASN
1	A	1422	LYS
1	A	2033	ARG
1	A	1848	ARG
1	A	1974	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides 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	2134/2410 (88%)	0.20	77 (3%) 42 32	94, 130, 186, 267	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	670	LEU	6.8
1	A	1315	ASN	6.2
1	A	548	VAL	5.2
1	A	2023	THR	4.8
1	A	1419	ARG	4.6

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 monosaccharides 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.