



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 11:35 PM UTC

PDB ID : 6OQ5 / pdb_00006oq5
Title : Structure of the full-length Clostridium difficile toxin B in complex with 3 VHHs
Authors : Chen, P.; Lam, K.; Jin, R.
Deposited on : 2019-04-25
Resolution : 3.87 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
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

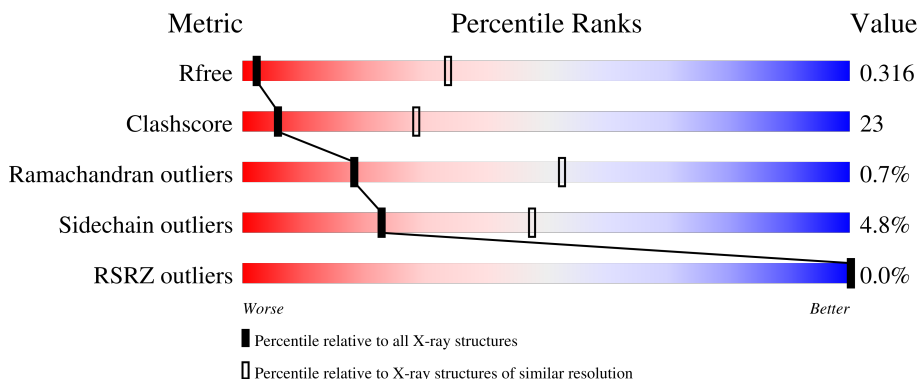
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.87 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	1186 (4.06-3.70)
Clashscore	190562	1012 (4.04-3.72)
Ramachandran outliers	187476	1167 (4.06-3.70)
Sidechain outliers	187428	1160 (4.06-3.70)
RSRZ outliers	180081	1185 (4.06-3.70)

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	2373	 61% 35% . .
2	D	153	 52% 24% 7% 18% %
3	E	137	 55% 23% . 22%
4	F	142	 51% 28% . 18%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 21503 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 Toxin B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	2346	Total	C	N	O	S	0	0	0
			18837	12009	2961	3820	47			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2368	HIS	-	expression tag	UNP M4NKV9
A	2369	HIS	-	expression tag	UNP M4NKV9
A	2370	HIS	-	expression tag	UNP M4NKV9
A	2371	HIS	-	expression tag	UNP M4NKV9
A	2372	HIS	-	expression tag	UNP M4NKV9
A	2373	HIS	-	expression tag	UNP M4NKV9

- Molecule 2 is a protein called 5D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	126	Total	C	N	O	S	0	0	0
			986	617	179	187	3			

- Molecule 3 is a protein called E3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	107	Total	C	N	O	S	0	0	0
			802	498	141	159	4			

- Molecule 4 is a protein called 7F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	116	Total	C	N	O	S	0	0	0
			876	544	154	172	6			

- Molecule 5 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by

depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Zn	0	0
			1	1		

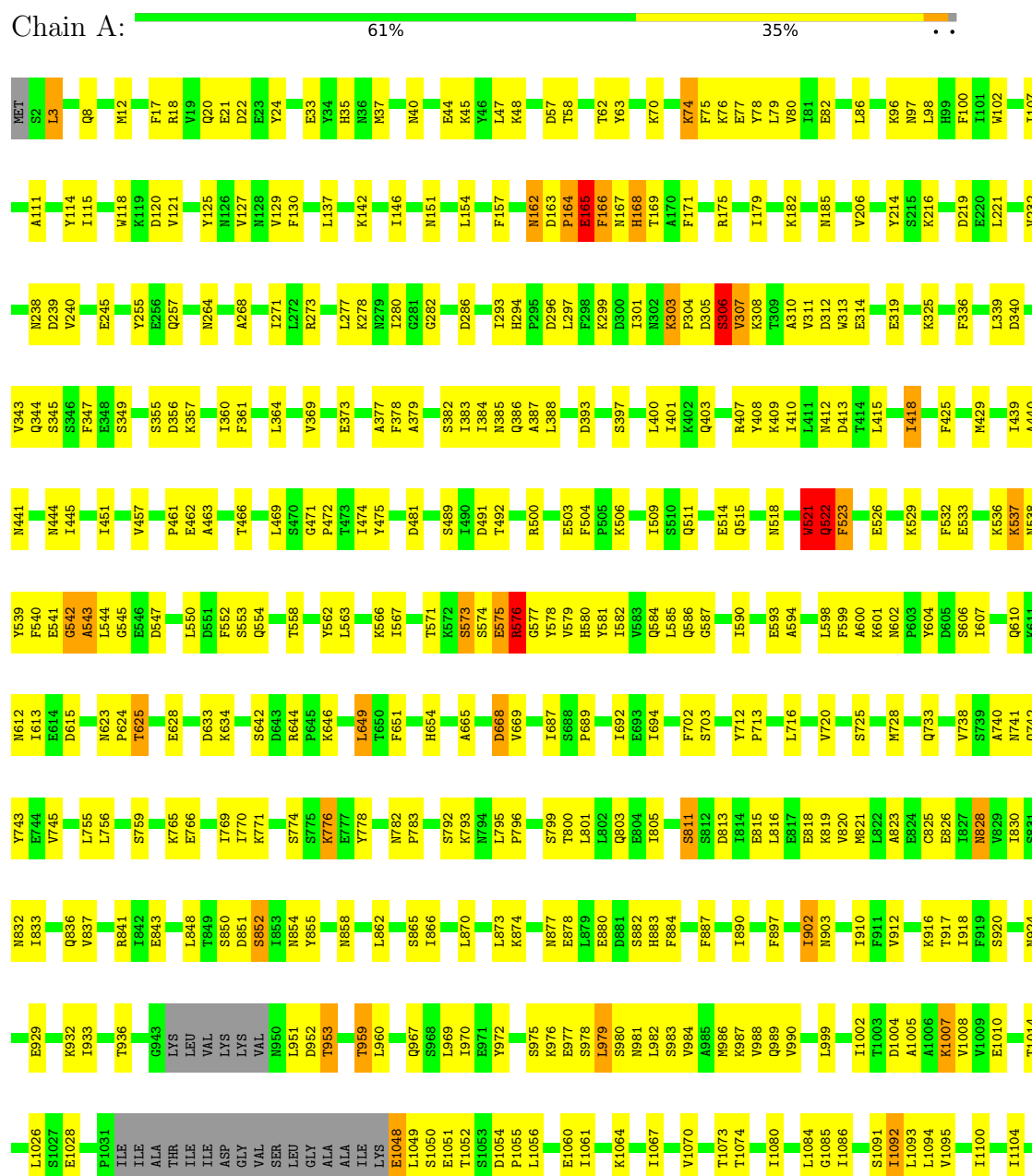
- Molecule 6 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Mg	0	0
			1	1		

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: Toxin B

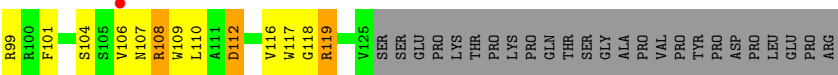


L1107	L1203	P1302	V1419	T1521	R1653	Y1741	G1830	L1915	Y1999	Y2110	D2215	Y2305
V1108	L1204	E1307	L1422	K1522	Q1654	V1742	M1831	N1918	F2000	Y2111	M2216	F2306
N1109	D1205	Y1308	L1430	N1526	M1655	W1743	M1832	A1923	V2005	F2112	E2219	Q2309
E1111	V1206	I1309	N1525	T1529	I1657		V1833	I1924	M2006	N2113	S2220	N2310
L1112	Q1209	L1311	N1440	G1530	N1661	I1752	G1835	L1930	Y2010	D2114	D2221	T2311
R1115	K1210	E1312	S1441	L1663	Y1662	S1754	L1836	L1938	F2018	G2116	K2222	W2315
K1120	L1216	L1313	N1442	L1664	D1665	S1785	I1837	D1933	F2019	M2118	Y2223	F2316
L1130	L1220	Y1315	H1443	S1314	L1666	S1786	I1838	A1940	F2020	Q2119	F2225	E2317
V1131	L1223	S1320	I1444	Y1531	L1667	E1757	M1839	E1942	A2021	W2120	D2226	G2321
E1134	L1227	G1321	Q1445	L1533	D1666	K1760	D1841	F1939	E2022	G2121	P2227	I2321
L1140	P1227	Y1331	K1447	L1545	G1668	Q1763	L1843	E1944	N2023	F2122	T2229	L2329
D1141	M1228	D1343	I1448	V1563	D1669	V1764	Y1844	N1942	E2024	I2125	K2231	D2330
D1142	R1229	D1344	F1453	I1566	S1671	R1767	F1846	N1943	E2025	V2129	K2233	F2336
K1143	A1232	W1345	I1461	S1579	T1673	F1768	K1847	R1944	Q2027	F2130	C2233	A2342
V1144	W1233	I1346	S1464	F1605	V1674		P1848	V1947	I2028	Y2131	K2234	
Q1148	W1237	V1349	E1472	M1610	S1678	N1784	P1849	E1948	G2029	F2132	G2235	V2347
D1149	T1238	D1350	Y1462	F1611	Q1679	F1785	V1850	W1949	T2030	G2136	I2236	G2347
L1151	P1239	N1351	S1464	I1612	K1680	S1786	N1852	K1950	F2031	I2137	N2237	F2356
V1152	G1240	V1352	E1472	T1613	V1681	G1787	L1853	E1951	K2032	I2138	E2238	
E1155	L1241	T1357	F1475	T1616	I1694	Q1788	T1854	G1954	D2033	V2142	D2241	Q2382
F1158	R1242	I1358	I1476	T1616	S1695	D1790	T1855	E1955	G2036	I2145	I2242	L2383
	E1245	K1362	N1477	T1616	P1696	V1793	G1856	M1956	F2037	D2153	Y2244	I2385
	T1249	I1363	T1480	T1616	I1698	I1796	F1857	F1959	K2038	I2157	F2246	H2389
	K1251	K1364	L1484	I1619	D1701	L1797	K1864	S1960	D2042	Y2171	M2252	HIS
G1166	L1251	K1365	F1485	Q1620	I1702	L1798	S1799	G1964	L2047	N2181	R2253	HIS
I1170	R1254	I1372	V1486	F1621	I1703	F1800	Y1865	F1967	E2051	I2182	T2254	HIS
W1171	L1255	L1373	S1487	E1623	T1706	T1801	F1866	L1970	G2052	Y2183	G2255	
R1172	L1255	S1374	E1488	F1624	P1707	P1802	F1867	L1971	E2053	Y2187	L2256	
M1173	E1262	T1375	L1489	I1625	V1708	S1803	P1869	Y1804	A2076	V2187	L2257	
E1174	W1265	I1378	P1490	C1626	Y1715	Y1805	T1870	Q1972	A2076		M2262	
S1177	K1265	E1379	D1491	D1627	I1719	E1806	Y1806	I1973	W2080	L2192	N2263	
G1178	F1268	K1382	V1493	E1628	I1719	D1807	S1876	G1974	D1975	Y2193	Y2264	
H1179	A1272	N1391	I1494	M1631	V1720	I1810	I1881	D1976	S2087	R2194	F2265	
T1180			I1495	T1632	N1724	L1814	I1882	K1977	K2088	Y2195	N2267	
I1185	L1275	V1396	K1497	Q1633	I1725	G1815	M1886	Y1978	Y2090	G2196	G2272	
D1186	L1276	N1397	Y1499	Y1635	N1727	L1816	F1889	N1980	A2098	D2198	Y2285	
H1187	K1280		K1504	F1639	E1728	S1818	V1894	N1981	Y2099	V2199	F2286	
F1188	Y1283	V1403	P1505	E1643	K1729	L1819	L1895	V1985	G2100	Y2201	V2291	
F1189		T1406	Y1510	Y1646	V1732	K1823	Q1896	M1986	L2102	F2202	M2292	
P1192	M1287	I1409	S1511	Y1647	I1734	F1824	S1901	Q1987	S2103	L2208	G2295	
T1194	S1298	A1415	M1512	L1648	N1735	I1825	F1906	N1994	I2105	E2209	T2310	
T1195	F1299	I1416	M1513	Y1649	I1826	I1826	F1906	D1995	N2106	G2211	G2211	
Y1196	I1300	K1515	K1514	V1650	S1738	N1827	K1907	N1996	D2107	Y2212	M2298	
E1197	V1301	I1417	K1515	G1651	I1739	M1828	Y1908	K1997	G2108	I2213	F2303	
E1198		E1418	D1516	M1652	R1740	F1829		H1998	Q2109	Y2214	K2304	

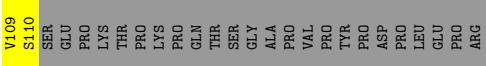
• Molecule 2: 5D



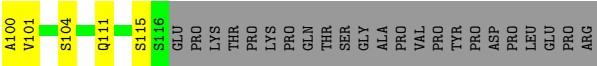
SER	G22	G23	A24	F27	T28	L29	Y32	G33	I34	G35	A36	F37	R45	Y50	I51	S52	A53	S54	A55	R56	T57	I58	L59	V64	F68	T69	I70	S71	R72	Y79	R83	R84	L86	Y94	T95	C96	A97	R98
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



● Molecule 3: E3



● Molecule 4: 7F



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	149.62Å 168.56Å 179.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.91 – 3.87 48.91 – 3.87	Depositor EDS
% Data completeness (in resolution range)	99.2 (48.91-3.87) 99.4 (48.91-3.87)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.37 (at 3.88Å)	Xtriage
Refinement program	REFMAC 5.8.0232	Depositor
R, R_{free}	0.263 , 0.315 0.263 , 0.316	Depositor DCC
R_{free} test set	2203 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	135.9	Xtriage
Anisotropy	0.169	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 92.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	21503	wwPDB-VP
Average B, all atoms (Å ²)	169.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.86% 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: ZN, MG

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	0.42	0/19208	1.02	29/26012 (0.1%)
2	D	0.46	0/1006	0.96	2/1360 (0.1%)
3	E	0.46	0/814	0.91	0/1098
4	F	0.48	0/893	0.92	0/1206
All	All	0.43	0/21921	1.01	31/29676 (0.1%)

There are no bond length outliers.

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2223	TYR	N-CA-C	9.68	124.39	112.59
1	A	1195	THR	CB-CA-C	7.61	123.56	109.86
1	A	1446	GLN	N-CA-C	-6.91	103.44	110.97
1	A	543	ALA	N-CA-C	-6.46	105.03	113.17
2	D	112	ASP	N-CA-C	-6.44	105.58	113.50

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	18837	0	18156	831	0
2	D	986	0	959	90	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	802	0	797	33	0
4	F	876	0	856	41	0
5	A	1	0	0	0	0
6	A	1	0	0	0	0
All	All	21503	0	20768	978	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 23.

The worst 5 of 978 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:2107:ASP:CB	1:A:2137:ILE:HG22	1.73	1.19
1:A:1847:LYS:HG3	1:A:1848:PRO:HD2	1.27	1.16
1:A:2025:GLU:O	1:A:2026:MET:HG2	1.44	1.14
1:A:120:ASP:HB3	1:A:357:LYS:HD2	1.18	1.14
1:A:625:THR:HB	1:A:628:GLU:HG2	1.25	1.14

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	2340/2373 (99%)	2001 (86%)	319 (14%)	20 (1%)	14	47
2	D	124/153 (81%)	112 (90%)	12 (10%)	0	100	100
3	E	105/137 (77%)	94 (90%)	11 (10%)	0	100	100
4	F	114/142 (80%)	98 (86%)	16 (14%)	0	100	100
All	All	2683/2805 (96%)	2305 (86%)	358 (13%)	20 (1%)	18	52

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	164	PRO
1	A	306	SER
1	A	307	VAL
1	A	574	SER
1	A	576	ARG

5.3.2 Protein sidechains ⓘ

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	2113/2142 (99%)	2017 (96%)	96 (4%)	24	49
2	D	102/127 (80%)	91 (89%)	11 (11%)	6	24
3	E	89/117 (76%)	87 (98%)	2 (2%)	45	64
4	F	95/119 (80%)	88 (93%)	7 (7%)	13	38
All	All	2399/2505 (96%)	2283 (95%)	116 (5%)	23	48

5 of 116 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1817	VAL
4	F	61	ASP
1	A	2025	GLU
4	F	57	ILE
2	D	52	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1210	GLN
1	A	2322	ASN
1	A	1645	ASN
1	A	2310	ASN
2	D	74	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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	2346/2373 (98%)	-0.55	0 100 100	80, 160, 233, 324	0
2	D	126/153 (82%)	-0.46	1 (0%) 82 63	147, 200, 255, 290	0
3	E	107/137 (78%)	-0.53	0 100 100	213, 260, 307, 341	0
4	F	116/142 (81%)	-0.72	0 100 100	117, 166, 214, 240	0
All	All	2695/2805 (96%)	-0.56	1 (0%) 100 100	80, 165, 249, 341	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	106	VAL	2.6

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.

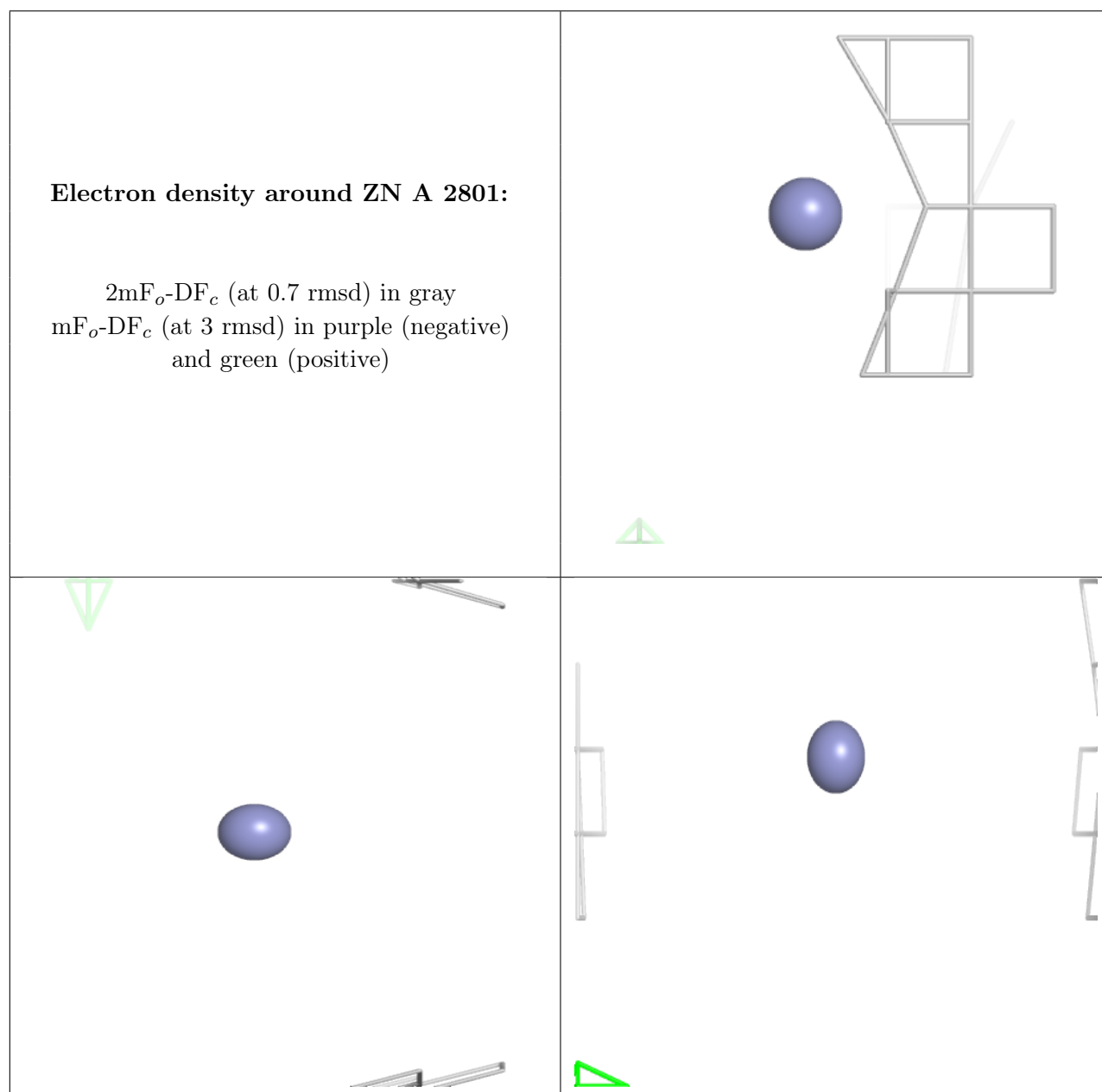
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	MG	A	2802	1/1	0.89	0.12	119,119,119,119	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
5	ZN	A	2801	1/1	1.00	0.02	120,120,120,120	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



6.5 Other polymers [i](#)

There are no such residues in this entry.