



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 20, 2024 – 10:47 AM EDT

PDB ID : 1EWR
Title : CRYSTAL STRUCTURE OF TAQ MUTS
Authors : Obmolova, G.; Ban, C.; Hsieh, P.; Yang, W.
Deposited on : 2000-04-26
Resolution : 3.19 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
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.20.1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.003 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

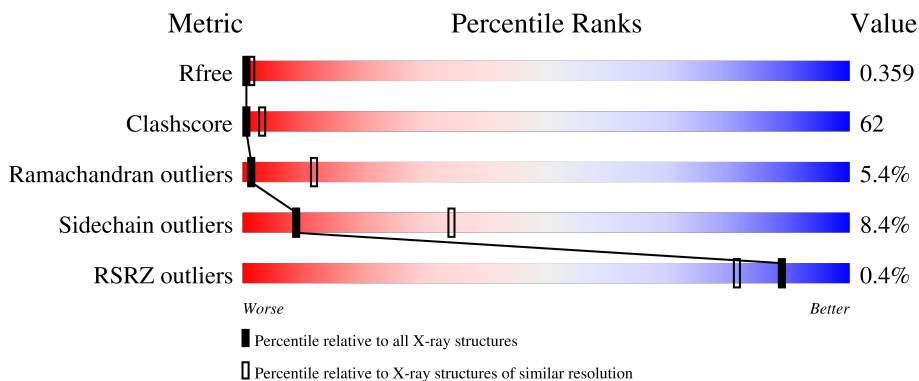
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.19 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}	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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	649	
1	B	649	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 7593 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 DNA MISMATCH REPAIR PROTEIN MUTS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	447	3468	2209	626	626	7	0	0	0
1	B	529	4125	2623	740	755	7	0	0	0

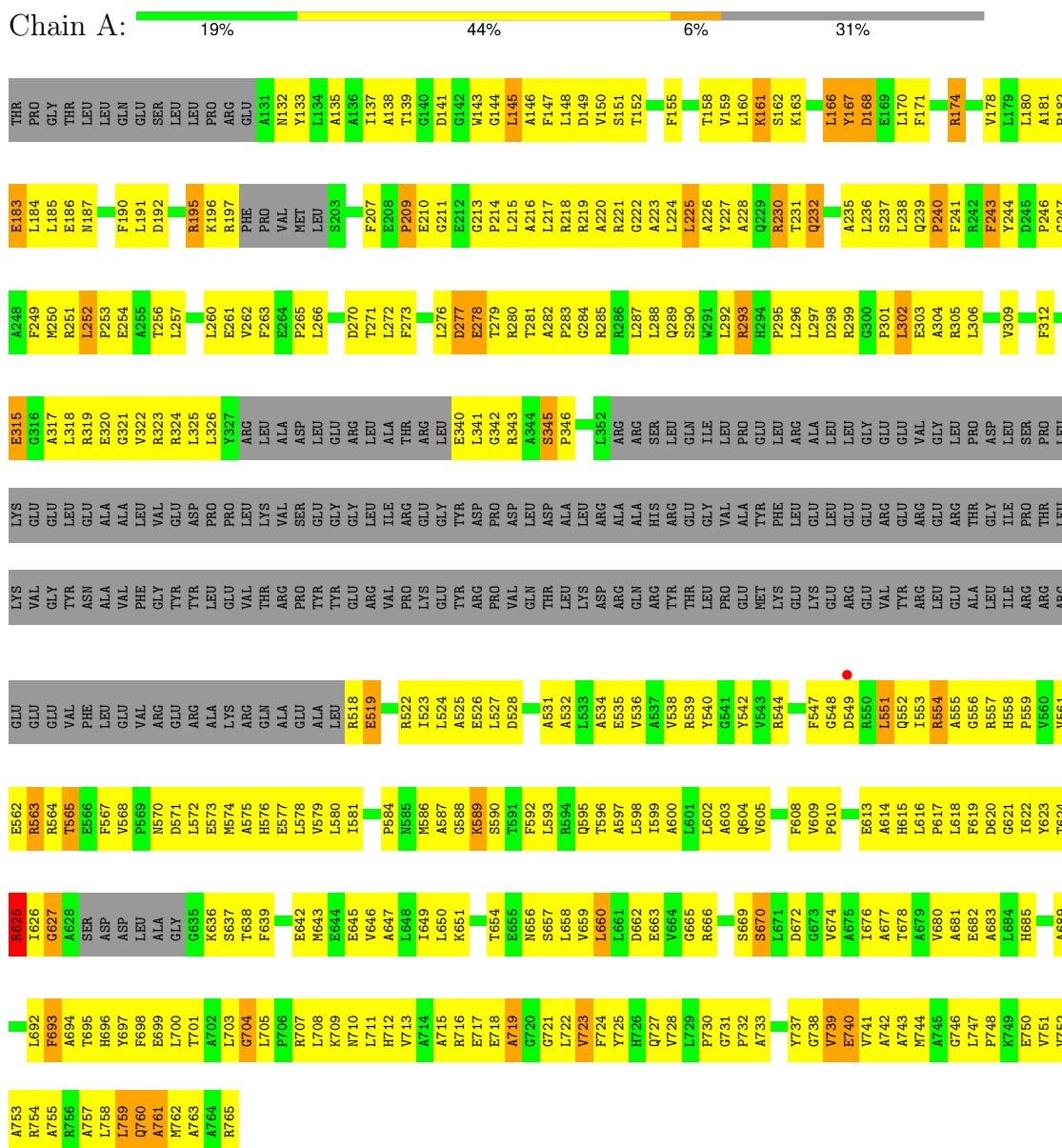
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	250	MSE	MET	modified residue	UNP Q56215
A	574	MSE	MET	modified residue	UNP Q56215
A	586	MSE	MET	modified residue	UNP Q56215
A	640	MSE	MET	modified residue	UNP Q56215
A	643	MSE	MET	modified residue	UNP Q56215
A	744	MSE	MET	modified residue	UNP Q56215
A	762	MSE	MET	modified residue	UNP Q56215
B	1250	MSE	MET	modified residue	UNP Q56215
B	1574	MSE	MET	modified residue	UNP Q56215
B	1586	MSE	MET	modified residue	UNP Q56215
B	1640	MSE	MET	modified residue	UNP Q56215
B	1643	MSE	MET	modified residue	UNP Q56215
B	1744	MSE	MET	modified residue	UNP Q56215
B	1762	MSE	MET	modified residue	UNP Q56215

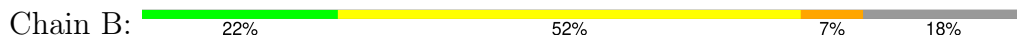
3 Residue-property plots

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: DNA MISMATCH REPAIR PROTEIN MUTS



- Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS



L1765	A1757	L1758	L1759	Q1760	A1761	M1762	ALA	ALA	ARG																																																								
L1692	F1693	A1694	T1695	H1696	V1697	F1698	E1699	L1700	T1701	A1702	L1703	G1704	L1705	R1706	R1707	L1708	K1709	M1710	L1711	H1712	A1713	V1714	A1715	A1716	E1717	E1718	A1719	G1720	G1721	L1722	V1723	F1724	Y1725	H1726	V1727	L1728	P1730	G1731	P1732	A1733	Y1737	G1738	V1739	E1740	A1741	A1742	A1743	M1744	L1747	P1748	K1749	E1750	V1751	V1752	A1753	R1754									
T1565	E1566	F1567	V1568	P1569	N1570	D1571	E1572	L1573	A1574	M1574	A1575	H1576	A1577	L1578	V1579	L1580	I1581	P1584	A1587	G1588	K1589	S1590	T1591	F1592	L1593	L1594	T1595	T1596	A1597	L1598	I1599	A1600	L1601	L1602	A1603	Q1604	V1605	F1608	Y1641	Y1642	V1643	P1610	E1613	A1614	H1615	L1616	P1617	L1618	F1619	A1620	G1621	L1622	Y1623	T1624	R1625	I1626	G1627	A1628							
SER	ASP	ASP	LEU	ALA	GLY	G1636	K1636	S1637	T1638	F1639	M1640	V1641	E1642	M1643	E1644	V1645	V1646	A1647	L1648	L1649	L1650	K1651	T1654	E1655	M1656	S1657	L1658	V1659	L1660	L1661	D1662	E1663	G1665	R1666	G1667	L1668	S1669	L1670	D1672	G1673	V1674	A1675	L1676	A1677	A1678	A1679	V1680	A1681	E1682	A1683	L1684	H1685	A1689	Y1690	V1691	T1691									
E1501	V1502	F1503	L1504	E1505	V1506	R1507	E1508	R1509	L1510	K1511	R1512	E1515	A1516	L1517	R1518	E1519	A1520	A1521	L1522	J1523	K1523	L1524	A1525	E1526	L1527	D1528	V1529	Y1530	A1531	A1532	L1533	A1534	E1535	V1536	L1537	V1538	R1539	Y1540	G1541	Y1542	V1543	R1544	P1545	F1547	L1551	Q1552	T1553	R1554	A1555	L1556	R1557	H1558	V1561	E1562	R1563	R1564									
VAL	GLY	TYR	ASN	ALA	VAL	PHE	GLY	TYR	TYR	LEU	GLU	VAL	THR	ARG	PRO	TYR	TYR	GLU	VAL	PRO	LYS	GLY	THR	ARG	PRO	VAL	VAL	THR	GLN	ARG	TYR	THR	LEU	LYS	PRO	LEU	PRO	GLU	MET	LYS	GLY	VAL	GLU	VAL	TYR	ARG	LEU	GLU	THR	ALA	LEU	ILE	ARG	THR	LEU	LYS	R1498								
E1380	E1381	L1382	E1383	A1384	A1385	L1386	V1387	E1388	L1389	P1390	P1391	LEU	LYS	VAL	SER	GLY	GLY	GLY	LEU	LEU	ILE	ASP	PRO	PRO	ASP	ASP	LEU	ASP	ASP	ASP	ALA	LEU	LEU	ARG	ALA	ALA	HIS	ARG	ARG	GLY	GLY	VAL	VAL	ALA	TYR	PHE	LYS	LEU	GLU	LEU	LEU	GLU	GLU	GLU	ARG	ARG	GLU	ARG	THR	THR	ILE	PRO	THR	LEU	LYS
G1247	M1250	A1251	L1252	P1253	E1254	A1255	T1256	L1257	R1258	E1261	V1262	F1263	L1266	R1267	G1268	Q1269	D1270	E1271	T1272	L1273	D1277	E1278	T1279	R1280	L1281	T1282	A1283	P1284	G1285	R1286	L1287	L1288	Q1289	S1290	W1291	L1292	L1293	R1294	H1295	P1296	L1297	D1298	L1299	R1300	P1301	G1302	E1303	G1304	A1305	R1306	V1309	F1312	V1313	R1314											
E1315	G1316	A1317	L1318	R1319	E1320	G1321	V1322	R1323	D1324	L1325	L1326	Y1327	R1328	D1331	E1332	E1333	R1334	L1335	A1336	T1337	R1338	R1339	L1340	L1341	G1342	R1343	A1344	S1345	P1346	K1347	D1348	A1351	L1352	R1353	R1354	S1355	L1356	Q1357	I1358	L1359	P1360	E1361	L1362	R1363	L1366	G1367	E1368	E1369	V1370	P1373	S1376	P1377	L1378	K1379											
L1180	A1181	P1182	E1183	L1184	L1185	E1186	N1187	F1190	L1191	D1192	F1193	F1194	R1195	M1196	ARG	PHE	PRO	VAL	MET	L1202	L1203	S1204	E1205	F1206	F1207	E1208	P1209	E1210	G1211	P1214	L1215	A1216	L1217	R1218	R1219	A1220	R1221	G1222	A1223	L1224	L1225	A1226	Y1227	A1228	T1231	Q1232	G1233	G1234	A1235	L1236	Q1239	Y1244	D1245	P1246											
T1117	P1118	G1119	T1120	L1121	L1122	Q1123	E1124	S1125	L1126	L1127	P1128	E1130	A1131	M1132	L1133	L1134	A1135	A1136	L1137	A1138	T1139	G1140	D1141	G1142	W1143	G1144	L1145	A1146	F1147	L1148	D1149	V1150	S1151	T1152	G1153	E1154	F1155	T1158	V1159	L1160	K1161	S1162	K1163	L1166	Y1167	D1168	E1169	L1170	F1171	R1172	H1173	R1174	E1177	V1178	L1179										

4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	96.73Å 96.73Å 427.13Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.94 – 3.19 19.94 – 3.20	Depositor EDS
% Data completeness (in resolution range)	93.0 (19.94-3.19) 94.3 (19.94-3.20)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 3.22Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.331 , 0.361 0.332 , 0.359	Depositor DCC
R_{free} test set	1813 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	98.6	Xtrriage
Anisotropy	0.224	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 155.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.058 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	7593	wwPDB-VP
Average B, all atoms (Å ²)	101.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% 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	0.48	0/3526	0.67	0/4757
1	B	0.51	0/4192	0.69	0/5662
All	All	0.50	0/7718	0.68	0/10419

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	3468	0	3513	435	0
1	B	4125	0	4200	526	0
All	All	7593	0	7713	949	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 62.

The worst 5 of 949 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:292:LEU:HD23	1:A:605:VAL:HG21	1.31	1.11
1:A:272:LEU:HD11	1:A:602:LEU:HD21	1.32	1.10

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:209:PRO:HB2	1:A:218:ARG:HH21	1.16	1.04
1:A:557:ARG:HH21	1:A:610:PRO:HA	1.26	1.00
1:B:1557:ARG:HH21	1:B:1610:PRO:HA	1.23	0.99

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	437/649 (67%)	344 (79%)	69 (16%)	24 (6%)	1	11
1	B	521/649 (80%)	406 (78%)	87 (17%)	28 (5%)	1	12
All	All	958/1298 (74%)	750 (78%)	156 (16%)	52 (5%)	1	12

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	187	ASN
1	A	235	ALA
1	A	519	GLU
1	A	670	SER
1	A	704	GLY

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	346/515 (67%)	316 (91%)	30 (9%)	8	32
1	B	419/515 (81%)	385 (92%)	34 (8%)	9	36
All	All	765/1030 (74%)	701 (92%)	64 (8%)	9	34

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1662	ASP
1	B	1693	PHE
1	A	670	SER
1	A	662	ASP
1	B	1739	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1604	GLN
1	B	1727	GLN
1	B	1685	HIS
1	B	1123	GLN
1	B	1585	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](#)

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	440/649 (67%)	-0.92	1 (0%) 92 87	34, 104, 154, 181	0
1	B	522/649 (80%)	-0.97	3 (0%) 85 76	14, 98, 158, 188	0
All	All	962/1298 (74%)	-0.95	4 (0%) 89 81	14, 102, 156, 188	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	549	ASP	2.3
1	B	1268	GLY	2.3
1	B	1142	GLY	2.1
1	B	1635	GLY	2.0

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.