W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

Full wwPDB X-ray Structure Validation Report (i)

Oct 16, 2021 – 09:09 PM EDT

PDB ID	:	1T2U
Title	:	Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1: structure of BRCA1 missense variant V1809F
Authors	:	Williams, R.S.; Lee, M.S.; Duong, D.D.; Glover, J.N.M.
Deposited on	:	2004-04-22
Resolution	:	2.80 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

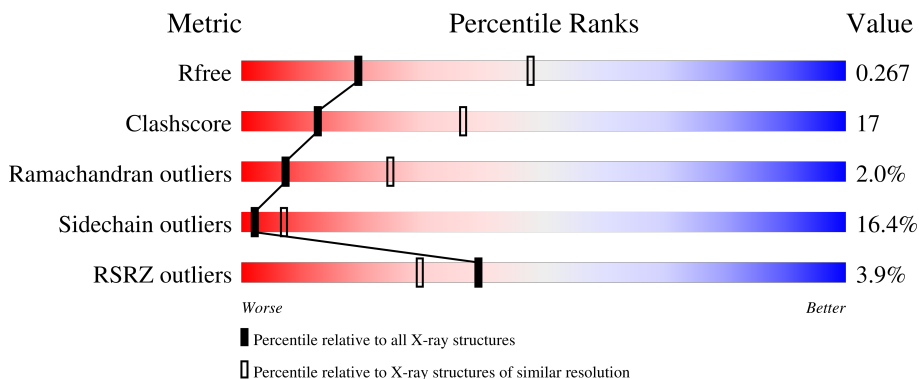|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.80 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	Percentile Ranks	Value
Rfree		0.267
Clashscore		17
Ramachandran outliers		2.0%
Sidechain outliers		16.4%
RSRZ outliers		3.9%

Worse ... *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 <=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	214	4% / 56% / 34% / 7% / ·

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 1634 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 Breast cancer type 1 susceptibility protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	207	Total	C	N	O	S	0	0	0
			1572	1011	254	293	14			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1809	PHE	VAL	engineered mutation	UNP P38398

- Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Co	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 5	O 4	S 1	0	0
3	A	1	Total 5	O 4	S 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	51	Total 51	O 51	0	0

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: Breast cancer type 1 susceptibility protein
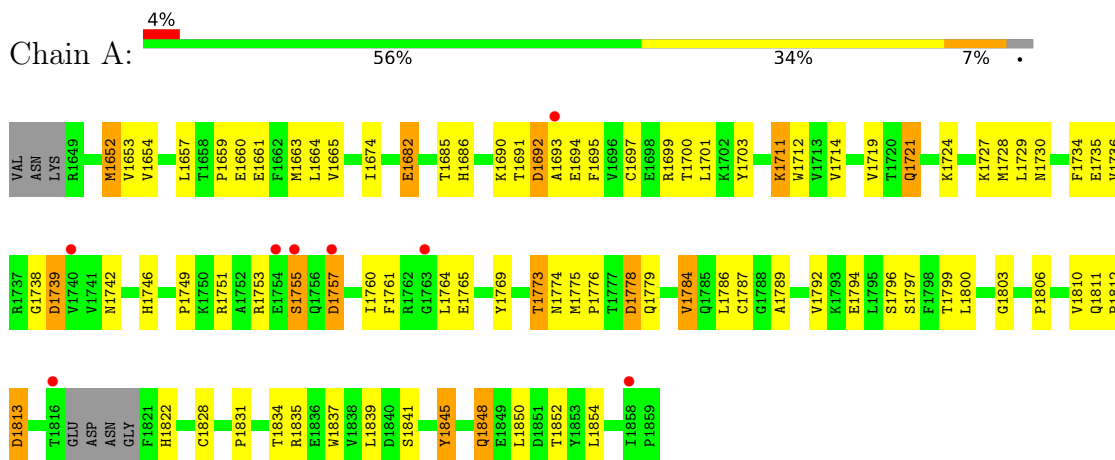
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	113.92Å 113.92Å 120.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.80 19.74 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.3 (20.00-2.80) 99.3 (19.74-2.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.41 (at 2.79Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.275 , 0.294 0.251 , 0.267	Depositor DCC
R_{free} test set	812 reflections (6.88%)	wwPDB-VP
Wilson B-factor (Å^2)	80.3	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 70.3	EDS
L-test for twinning[2]	$< |L| > = 0.41, < L^2 > = 0.23$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	1634	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.37% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CO, SO4

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	0.84	3/1610 (0.2%)	0.91	3/2196 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1692	ASP	CB-CG	9.70	1.72	1.51
1	A	1692	ASP	CG-OD2	7.07	1.41	1.25
1	A	1692	ASP	C-O	5.15	1.33	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1692	ASP	CB-CG-OD2	12.86	129.87	118.30
1	A	1778	ASP	CB-CG-OD2	6.61	124.25	118.30
1	A	1739	ASP	CB-CG-OD2	6.34	124.00	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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	1572	0	1476	51	0
2	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	10	0	0	0	0
4	A	51	0	0	4	0
All	All	1634	0	1476	51	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 17.

All (51) 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:1850:LEU:O	1:A:1854:LEU:HG	1.86	0.74
1:A:1811:GLN:HB3	4:A:14:HOH:O	1.91	0.71
1:A:1828:CYS:HA	4:A:19:HOH:O	1.93	0.67
1:A:1806:PRO:O	1:A:1831:PRO:HD2	1.97	0.64
1:A:1761:PHE:HB3	1:A:1789:ALA:HB2	1.83	0.60
1:A:1813:ASP:OD1	1:A:1813:ASP:N	2.36	0.58
1:A:1721:GLN:OE1	1:A:1724:LYS:HE2	2.04	0.58
1:A:1760:ILE:HG21	1:A:1845:TYR:O	2.05	0.57
1:A:1757:ASP:N	1:A:1757:ASP:OD1	2.39	0.55
1:A:1760:ILE:HG12	1:A:1845:TYR:CE1	2.42	0.54
1:A:1784:VAL:HG13	1:A:1789:ALA:HB3	1.88	0.54
1:A:1700:THR:O	1:A:1703:TYR:HB3	2.08	0.53
1:A:1773:THR:OG1	1:A:1774:ASN:N	2.41	0.53
1:A:1711:LYS:HE3	4:A:38:HOH:O	2.07	0.53
1:A:1835:ARG:HD3	1:A:1839:LEU:HD11	1.91	0.51
1:A:1692:ASP:O	1:A:1692:ASP:OD1	2.30	0.50
1:A:1652:MET:HG2	1:A:1686:HIS:HB2	1.92	0.50
1:A:1739:ASP:OD1	1:A:1742:ASN:N	2.36	0.49
1:A:1749:PRO:O	1:A:1753:ARG:HG3	2.12	0.48
1:A:1792:VAL:HG13	1:A:1797:SER:HB2	1.95	0.48
1:A:1765:GLU:HG2	1:A:1792:VAL:HG23	1.95	0.48
1:A:1660:GLU:H	1:A:1660:GLU:CD	2.17	0.48
1:A:1738:GLY:CA	1:A:1746:HIS:CE1	2.97	0.48
1:A:1665:VAL:HG22	1:A:1719:VAL:HG21	1.95	0.47
1:A:1685:THR:O	1:A:1711:LYS:HB3	2.15	0.47
1:A:1659:PRO:O	1:A:1663:MET:HG2	2.15	0.47
1:A:1711:LYS:HB3	1:A:1712:TRP:H	1.56	0.46
1:A:1776:PRO:HG2	1:A:1779:GLN:HG2	1.97	0.46
1:A:1695:PHE:CD2	1:A:1695:PHE:N	2.84	0.46
1:A:1701:LEU:HD13	1:A:1775:MET:HG2	1.98	0.46
1:A:1811:GLN:HA	1:A:1812:PRO:HD3	1.73	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1837:TRP:O	1:A:1841:SER:OG	2.14	0.45
1:A:1761:PHE:HA	1:A:1764:LEU:HD12	1.98	0.45
1:A:1657:LEU:HB3	1:A:1661:GLU:HB3	1.99	0.44
1:A:1697:CYS:O	1:A:1739:ASP:HA	2.17	0.44
1:A:1761:PHE:HB2	1:A:1787:CYS:O	2.17	0.43
1:A:1691:THR:HG22	1:A:1697:CYS:HB3	2.00	0.43
1:A:1835:ARG:HD3	1:A:1839:LEU:CD1	2.49	0.43
1:A:1695:PHE:CE1	1:A:1734:PHE:HA	2.54	0.42
1:A:1841:SER:HA	1:A:1848:GLN:HE21	1.84	0.42
1:A:1714:VAL:HA	1:A:1735:GLU:HA	2.01	0.42
1:A:1751:ARG:O	1:A:1755:SER:OG	2.33	0.42
1:A:1692:ASP:OD1	1:A:1692:ASP:C	2.58	0.42
1:A:1761:PHE:HD1	1:A:1787:CYS:HB2	1.85	0.42
1:A:1769:TYR:HB3	1:A:1810:VAL:HG22	2.02	0.41
1:A:1682:GLU:CD	1:A:1682:GLU:H	2.22	0.41
1:A:1839:LEU:HD23	1:A:1839:LEU:HA	1.83	0.41
1:A:1652:MET:HG3	1:A:1674:ILE:HG21	2.01	0.41
1:A:1738:GLY:HA3	1:A:1746:HIS:CE1	2.56	0.41
1:A:1692:ASP:O	1:A:1694:GLU:N	2.54	0.41
1:A:1796:SER:HB3	4:A:2:HOH:O	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	203/214 (95%)	186 (92%)	13 (6%)	4 (2%)	7	24

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1693	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1799	THR
1	A	1800	LEU
1	A	1803	GLY

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	165/191 (86%)	138 (84%)	27 (16%)	2	7

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1652	MET
1	A	1653	VAL
1	A	1654	VAL
1	A	1664	LEU
1	A	1682	GLU
1	A	1690	LYS
1	A	1699	ARG
1	A	1711	LYS
1	A	1721	GLN
1	A	1727	LYS
1	A	1728	MET
1	A	1729	LEU
1	A	1730	ASN
1	A	1736	VAL
1	A	1755	SER
1	A	1757	ASP
1	A	1773	THR
1	A	1778	ASP
1	A	1784	VAL
1	A	1786	LEU
1	A	1794	GLU
1	A	1813	ASP
1	A	1822	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1834	THR
1	A	1845	TYR
1	A	1848	GLN
1	A	1852	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1742	ASN
1	A	1746	HIS
1	A	1785	GLN
1	A	1826	GLN
1	A	1848	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	$\#|Z| > 2$
3	SO4	A	201	-	4,4,4	0.11	0	6,6,6	0.41	0
3	SO4	A	202	-	4,4,4	0.29	0	6,6,6	0.33	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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	207/214 (96%)	-0.05	8 (3%)	39	29	9, 28, 37, 42	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1757	ASP	3.2
1	A	1754	GLU	2.8
1	A	1740	VAL	2.6
1	A	1693	ALA	2.5
1	A	1816	THR	2.5
1	A	1858	ILE	2.5
1	A	1755	SER	2.5
1	A	1763	GLY	2.2

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
3	SO4	A	202	5/5	0.90	0.24	65,66,66,66	0
3	SO4	A	201	5/5	0.93	0.12	77,78,78,78	0
2	CO	A	101	1/1	0.95	0.23	43,43,43,43	0

6.5 Other polymers (i)

There are no such residues in this entry.