Full wwPDB X-ray Structure Validation Report (i)

Jan 5, 2026 – 04:57 pm GMT

PDB ID	:	9QPX / pdb_00009qpx
Title	:	BRCA1 BRCT tandem repeat with RNA polymerase II pSer5-CTD peptide
Authors	:	Houser, J.; Klapstova, V.; Sebesta, M.
Deposited on	:	2025-03-30
Resolution	:	3.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (1)) were used in the production of this report:

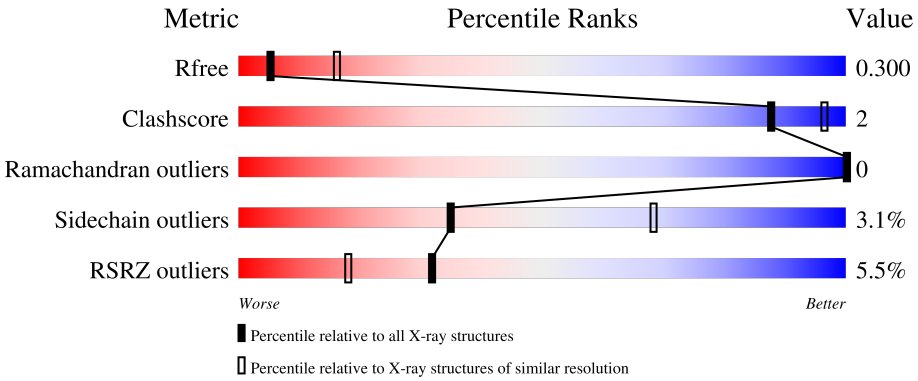
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	?.? (???), CSD ??CSD?? (????)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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.



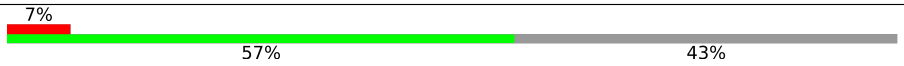
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	2511 (3.00-3.00)
Clashscore	180529	2866 (3.00-3.00)
Ramachandran outliers	177936	2778 (3.00-3.00)
Sidechain outliers	177891	2781 (3.00-3.00)
RSRZ outliers	164620	2523 (3.00-3.00)

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	215	7% 92% 7% ·
1	B	215	6% 92% 6% ·
1	C	215	3% 88% 9% ••
1	D	215	6% 89% 8% ••
2	E	14	7% 71% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	14	
2	G	14	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6819 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	211	Total	C	N	O	S	0	0	0
			1665	1067	276	308	14			
1	D	212	Total	C	N	O	S	0	0	0
			1686	1078	284	310	14			
1	B	211	Total	C	N	O	S	0	0	0
			1651	1057	273	307	14			
1	C	211	Total	C	N	O	S	0	0	0
			1607	1033	264	296	14			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1645	ALA	-	expression tag	UNP P38398
D	1645	ALA	-	expression tag	UNP P38398
B	1645	ALA	-	expression tag	UNP P38398
C	1645	ALA	-	expression tag	UNP P38398

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB1.

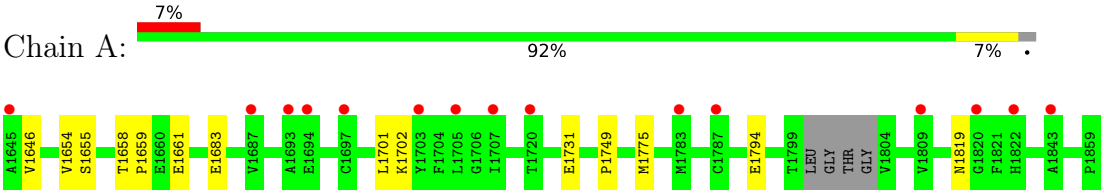
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	9	Total	C	N	O	P	0	0	0
			69	41	9	18	1			
2	E	10	Total	C	N	O	P	0	0	0
			79	44	10	23	2			
2	G	8	Total	C	N	O	P	0	0	0
			62	37	8	16	1			

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.

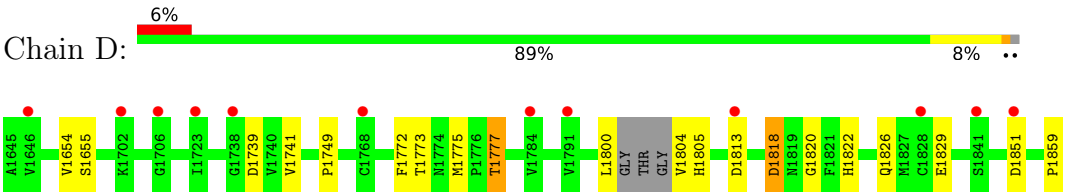
• Molecule 1: Breast cancer type 1 susceptibility protein

Chain A:



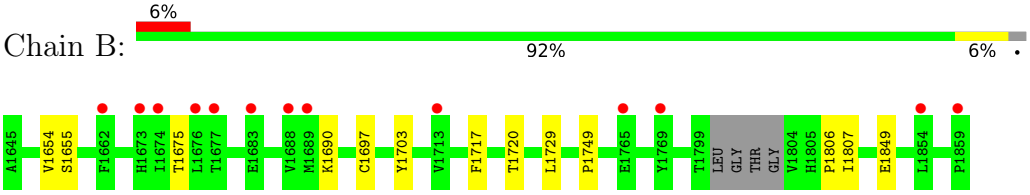
• Molecule 1: Breast cancer type 1 susceptibility protein

Chain D:



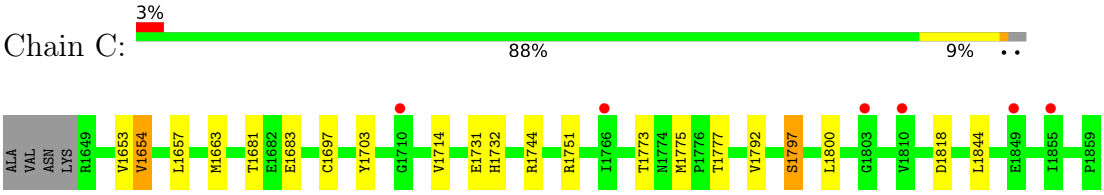
• Molecule 1: Breast cancer type 1 susceptibility protein

Chain B:



• Molecule 1: Breast cancer type 1 susceptibility protein

Chain C:

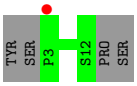


• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain F:

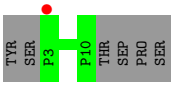
• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain E:



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain G:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	112.83Å 134.76Å 180.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.37 – 3.00 49.37 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.37-3.00) 99.7 (49.37-3.00)	Depositor EDS
R_{merge}	0.29	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.39 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.255 , 0.312 0.265 , 0.300	Depositor DCC
R_{free} test set	1386 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	76.3	Xtriage
Anisotropy	0.586	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 60.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6819	wwPDB-VP
Average B, all atoms (Å^2)	103.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.43% 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: SEP

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	0.49	0/1705	1.02	0/2317
1	B	0.47	0/1691	1.00	0/2302
1	C	0.49	0/1648	1.01	3/2251 (0.1%)
1	D	0.49	0/1726	1.03	2/2344 (0.1%)
2	E	0.54	0/61	0.81	0/82
2	F	0.57	0/61	1.09	0/82
2	G	0.47	0/54	0.97	0/72
All	All	0.49	0/6946	1.02	5/9450 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	1818	ASP	CB-CA-C	-5.50	100.57	110.03
1	C	1818	ASP	CA-CB-CG	5.19	117.79	112.60
1	C	1732	HIS	CA-CB-CG	5.09	118.89	113.80
1	D	1777	THR	CA-CB-OG1	-5.06	102.01	109.60
1	C	1777	THR	CA-CB-OG1	-5.01	102.08	109.60

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	1665	0	1609	7	0
1	B	1651	0	1577	5	0
1	C	1607	0	1511	8	0
1	D	1686	0	1642	7	0
2	E	79	0	61	0	0
2	F	69	0	58	1	0
2	G	62	0	51	0	0
All	All	6819	0	6509	25	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 2.

All (25) 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:D:1818:ASP:HB3	1:D:1820:GLY:H	1.66	0.60
1:D:1804:VAL:HG22	1:D:1805:HIS:H	1.67	0.57
1:A:1701:LEU:HD13	1:A:1775:MET:SD	2.47	0.55
1:A:1658:THR:HG22	1:A:1661:GLU:HG3	1.90	0.52
1:D:1772:PHE:CD1	1:D:1777:THR:HG22	2.47	0.50
1:C:1751:ARG:NH2	1:C:1844:LEU:O	2.43	0.47
1:A:1731:GLU:OE1	1:D:1822:HIS:NE2	2.46	0.47
1:C:1654:VAL:HG13	1:C:1657:LEU:HD22	1.98	0.46
1:C:1792:VAL:CG1	1:C:1797:SER:O	2.63	0.46
1:B:1655:SER:HA	2:F:5:SEP:O2P	2.18	0.44
1:A:1658:THR:HG22	1:A:1661:GLU:CG	2.47	0.44
1:C:1714:VAL:HG11	1:C:1731:GLU:HB3	1.99	0.44
1:A:1658:THR:OG1	1:A:1659:PRO:HD2	2.18	0.43
1:A:1819:ASN:HB3	1:B:1729:LEU:O	2.19	0.43
1:B:1806:PRO:O	1:B:1807:ILE:HG13	2.18	0.42
1:D:1739:ASP:C	1:D:1741:VAL:H	2.28	0.42
1:C:1773:THR:O	1:C:1775:MET:N	2.52	0.42
1:B:1717:PHE:HA	1:B:1720:THR:OG1	2.19	0.41
1:A:1655:SER:HB2	1:A:1702:LYS:HD2	2.02	0.41
1:C:1654:VAL:CG1	1:C:1657:LEU:HD22	2.49	0.41
1:C:1697:CYS:SG	1:C:1703:TYR:CE1	3.14	0.41
1:D:1773:THR:O	1:D:1775:MET:N	2.54	0.40
1:D:1826:GLN:HG2	1:D:1859:PRO:O	2.21	0.40
1:B:1697:CYS:SG	1:B:1703:TYR:CE1	3.14	0.40
1:C:1681:THR:HG23	1:C:1683:GLU:H	1.87	0.40

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	207/215 (96%)	194 (94%)	13 (6%)	0	100	100
1	B	207/215 (96%)	195 (94%)	12 (6%)	0	100	100
1	C	209/215 (97%)	192 (92%)	17 (8%)	0	100	100
1	D	208/215 (97%)	196 (94%)	12 (6%)	0	100	100
2	E	7/14 (50%)	6 (86%)	1 (14%)	0	100	100
2	F	6/14 (43%)	4 (67%)	2 (33%)	0	100	100
2	G	5/14 (36%)	4 (80%)	1 (20%)	0	100	100
All	All	849/902 (94%)	791 (93%)	58 (7%)	0	100	100

There are no Ramachandran outliers to report.

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	181/191 (95%)	176 (97%)	5 (3%)	38	70
1	B	178/191 (93%)	173 (97%)	5 (3%)	38	70
1	C	168/191 (88%)	162 (96%)	6 (4%)	30	64
1	D	185/191 (97%)	178 (96%)	7 (4%)	28	62
2	E	8/12 (67%)	8 (100%)	0	100	100
2	F	8/12 (67%)	8 (100%)	0	100	100
2	G	7/12 (58%)	7 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	735/800 (92%)	712 (97%)	23 (3%)	35 68

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1646	VAL
1	A	1654	VAL
1	A	1683	GLU
1	A	1749	PRO
1	A	1794	GLU
1	D	1654	VAL
1	D	1655	SER
1	D	1749	PRO
1	D	1800	LEU
1	D	1813	ASP
1	D	1829	GLU
1	D	1851	ASP
1	B	1654	VAL
1	B	1675	THR
1	B	1690	LYS
1	B	1749	PRO
1	B	1849	GLU
1	C	1653	VAL
1	C	1654	VAL
1	C	1663	MET
1	C	1744	ARG
1	C	1797	SER
1	C	1800	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1732	HIS
1	D	1779	GLN
1	D	1848	GLN
1	B	1756	GLN
1	C	1779	GLN
1	C	1848	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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.

5.5 Carbohydrates

There are no oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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, 95^{th} 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	211/215 (98%)	0.51	15 (7%) 23 13	56, 91, 126, 178	0
1	B	211/215 (98%)	0.49	13 (6%) 28 15	72, 99, 132, 165	0
1	C	211/215 (98%)	0.27	6 (2%) 55 33	82, 120, 157, 188	0
1	D	212/215 (98%)	0.47	12 (5%) 30 17	64, 91, 128, 148	0
2	E	8/14 (57%)	0.27	1 (12%) 9 5	89, 111, 120, 140	0
2	F	8/14 (57%)	0.57	0 100 100	93, 114, 133, 141	0
2	G	7/14 (50%)	0.40	1 (14%) 7 4	102, 121, 134, 143	0
All	All	868/902 (96%)	0.43	48 (5%) 32 18	56, 101, 139, 188	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1768	CYS	4.1
1	A	1703	TYR	4.1
1	A	1822	HIS	3.6
1	B	1676	LEU	3.5
1	B	1859	PRO	3.5
1	B	1683	GLU	3.5
1	A	1697	CYS	3.3
1	B	1673	HIS	3.2
1	D	1841	SER	3.2
1	B	1677	THR	3.1
1	C	1855	ILE	3.1
1	A	1707	ILE	3.0
1	D	1706	GLY	3.0
1	B	1674	ILE	2.7
1	A	1705	LEU	2.7
1	A	1787	CYS	2.7
1	B	1662	PHE	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1783	MET	2.6
1	B	1713	VAL	2.5
1	D	1784	VAL	2.5
1	D	1791	VAL	2.5
1	C	1803	GLY	2.5
1	A	1720	THR	2.4
2	G	3	PRO	2.4
1	A	1694	GLU	2.3
1	D	1828	CYS	2.3
1	A	1645	ALA	2.3
1	A	1820	GLY	2.3
1	C	1810	VAL	2.3
1	A	1693	ALA	2.3
1	B	1689	MET	2.3
1	A	1687	VAL	2.2
1	D	1738	GLY	2.2
1	D	1702	LYS	2.2
1	C	1849	GLU	2.2
1	D	1813	ASP	2.2
1	B	1854	LEU	2.2
1	D	1646	VAL	2.2
1	C	1710	GLY	2.1
1	D	1723	ILE	2.1
1	D	1851	ASP	2.1
2	E	3	PRO	2.1
1	A	1809	VAL	2.1
1	B	1765	GLU	2.1
1	B	1769	TYR	2.1
1	C	1766	ILE	2.0
1	A	1843	ALA	2.0
1	B	1688	VAL	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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, 95^{th} 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
2	SEP	E	12	10/11	0.34	0.12	143,195,210,214	0
2	SEP	F	5	10/11	0.86	0.09	86,100,106,111	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SEP	G	5	10/11	0.90	0.07	91,98,108,126	0
2	SEP	E	5	10/11	0.92	0.09	70,75,82,84	0

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.