



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 1, 2017 – 06:34 PM EDT

PDB ID : 1MJE
Title : STRUCTURE OF A BRCA2-DSS1-SSDNA COMPLEX
Authors : Yang, H.; Jeffrey, P.D.; Miller, J.; Kinnucan, E.; Sun, Y.; Thoma, N.H.; Zheng, N.; Chen, P.L.; Lee, W.H.; Pavletich, N.P.
Deposited on : unknown
Resolution : 3.50 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

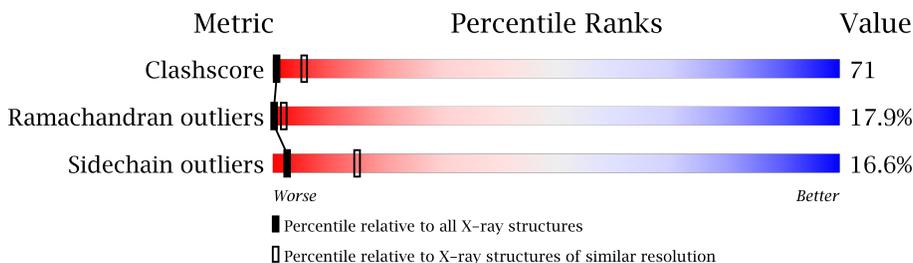
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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))
Clashscore	112137	1322 (3.60-3.40)
Ramachandran outliers	110173	1283 (3.60-3.40)
Sidechain outliers	110143	1284 (3.60-3.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	6	
2	B	70	
3	A	649	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5198 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 DNA chain called 5'-D(P*TP*TP*TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	6	121	60	12	43	6	16	0	0

- Molecule 2 is a protein called Deleted in split hand/split foot protein 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	42	335	209	50	76	0	0	0

- Molecule 3 is a protein called breast cancer 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	600	4742	3025	824	877	16	0	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

Note EDS was not executed.

- Molecule 1: 5'-D(P*TP*TP*TP*TP*TP*T)-3'

Chain C:  83% 17%

TS01
TS02
TS03
TS04
TS05
TS06

- Molecule 2: Deleted in split hand/split foot protein 1

Chain B:  17% 30% 10% 40%

NET SER GLU LYS LYS LYS ASP F7 V8 D9 L10 G11 L12 E15 D16 D17 E18 F19 E20 E21 E22 P23 A24 E25 ASP TRP ALA ALA LEU LEU ASP GLU ASP ASP ASP ASP ASP H37 V38 W39 W39 E40 D41 N42 W43 D44 D45 ASP ASN VAL VAL GLU D50 N54 Q55 L56 R57 A58 E59 L60 E61 K62 H63

GLY TYR MET MET GLU THR SER

- Molecule 3: breast cancer 2

Chain A:  20% 50% 20% 8%

ASN GLN SER LYS THR ASP GLY ASP ARG ARG HIS VAL ARG GLN PHE ASN K2399 D2400 A24 P23 E25 ASP TRP ALA ALA LEU LEU ASP GLU ASP ASP ASP ASP H37 V38 W39 W39 E40 D41 N42 W43 D44 D45 ASP ASN VAL VAL GLU D50 N54 Q55 L56 R57 A58 E59 L60 E61 K62 H63

I2442 S2443 L2444 Q2445 A2446 A2447 V2448 G2449 D2450 R2451 A2452 G2453 S2454 A2455 C2456 S2457 P2458 K2459 Q2460 L2461 Y2462 I2463 Y2464 G2465 V2466 S2467 K2468 E2469 C2470 I2471 S2472 N2473 V2474 N2475 S2476 K2477 N2478 E2479 Y2480 F2481 Q2482 F2483 D2484 R2485 Q2486 F2489 G2490 K2491 E2492 D2493 L2494 C2495 F2501 L2502 A2503

D2504 G2505 W2507 L2508 I2509 P2510 S2511 R2512 D2513 G2514 L2515 K2516 A2517 K2518 E2519 E2520 F2521 Q2522 R2523 A2524 L2525 C2526 D2527 T2528 P2529 V2530 G2531 D2532 P2533 K2534 L2535 S2537 I2539 W2540 V2541 A2542 N2543 H2544 Y2545 R2546 W2547 I2548 V2549 W2550 K2551 L2552 A2553 E2554 M2555 E2556 F2557 A2558 F2559 P2560 K2561 E2562 F2563

A2564 M2565 R2566 C2567 L2568 N2569 P2570 E2571 R2572 V2573 A2574 L2575 Q2576 L2577 K2578 Y2579 R2580 Y2581 D2582 V2583 I2585 A2584 D2586 R2589 R2590 S2591 Q2592 L2593 K2594 K2595 L2596 L2597 E2598 S2599 D2600 D2601 T2602 A2603 K2604 K2605 T2606 L2607 V2608 L2609 C2610 I2611 S2612 D2613 I2614 ILE SER PRU SER THR LYS VAL SER GLU THR

SER GLY LYS THR SER GLY ASP LYS V2638 D2638 T2639 I2640 E2641 L2642 T2643 D2644 G2645 W2646 Y2647 A2648 V2649 R2650 A2651 Q2652 L2653 D2654 P2655 P2656 L2657 L2660 S2663 G2664 K2665 L2666 T2667 V2668 G2669 Q2670 K2671 I2672 I2673 T2674 Q2675 G2676 A2677 E2678 V2679 G2680 G2681 S2682 P2683 D2684 A2685 C2686

K3088	L3108	E3030	L2964	Y3887	L3749	A3687
L3099	K3109	D3031	L2965	K2898	Q2750	P2688
I3100	W3110	I3032	Q2966	K2899	W2751	L2689
K3033	SER	K3033	V2967	K2900	V2752	E2690
P3034	THR	P3034	Q2968	E2901	E2753	A2691
R3035	PRO	R3035	Q2969	K2902	K2754	P2692
V3102	ASN	V3036	P2970	S2903	T2755	D2693
L3103		L3037	R2971	A2904	V2756	S2694
P3108		I3038		L2905	S2757	L2695
K3109		L2974		L2906	G2758	R2696
SER		E2975		S2907	L2759	L2697
THR		F2976		I2908	Y2760	K2698
PRO		S2977		R2909	Y2761	I2699
ASN		R2978		R2910	F2762	S2700
		L2979		P2911	R2763	A2701
		S2980		S2912	S2764	M2702
		D2981		S2913	E2765	S2703
		P2982		D2914	R2766	T2704
		A2983		L2915	E2767	R2705
		F2984			E2768	R2706
		Q2985		L2919	E2769	A2707
		P2986		G2922	K2770	R2708
		P2987		K2923	E2771	W2709
		C2988		R2924	A2772	H2710
		S2989		Y2925	L2773	S2711
		E2990		R2926	R2774	R2712
		V2991		I2927	F2775	L2713
		F3057		Y2928	A2776	G2714
		A3058		H2929	Q2779	F2715
		C3059		L2930	Q2780	F2716
				A2931	K2781	R2717
		S3062		V2932	K2782	P2721
		I3063		S2933	L2783	F2722
		E3070		K2934	E2784	P2723
		A3071		S2935	A2785	L2724
		Y3072		K2936	L2786	P2725
		F3073		S2937	F2787	L2726
		Q3074		R2938	T2788	S2727
		E3075		F2939	K2789	S2728
		K3076		E2940	V2790	L2729
		V3077		R2941	H2791	F2730
		N3078		P2942	T2792	S2731
		N3079		L2946	GLU	D2732
		L3080		T2947	GLY	G2733
		K3081			L2882	G2734
		H3082		R2951	S2883	N2735
		A3083		T2952	R2884	V2736
		E3084		Y2953	D2885	G2737
		E3085		Y2954	V2886	C2738
		N3086		Q2955	T2887	V2739
		I3087		Q2956	T2888	D2740
		D3089		Q2957	V2889	I2741
		T3090		L3020	W2890	I2742
		F3091		V3021	K2891	V2743
		Y3092		F3024	L2892	Q2744
		K3093			R2893	R2745
		E3096		L3028	V2894	V2746
		K3097		N3029	E2962	Y2747
					S2896	P2748

4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α , β , γ	198.87Å 198.87Å 200.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.50	Depositor
% Data completeness (in resolution range)	(Not available) (15.00-3.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.245 , 0.278	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	5198	wwPDB-VP
Average B, all atoms (Å ²)	105.0	wwPDB-VP

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	C	0.96	1/132 (0.8%)	0.85	0/200
2	B	0.66	0/340	0.88	1/460 (0.2%)
3	A	0.66	2/4843 (0.0%)	0.92	10/6557 (0.2%)
All	All	0.67	3/5315 (0.1%)	0.92	11/7217 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	A	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2790	VAL	CA-CB	7.32	1.70	1.54
1	C	501	DT	OP3-P	-6.11	1.53	1.61
3	A	2789	LYS	CG-CD	5.09	1.69	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	A	3005	GLY	N-CA-C	-8.10	92.84	113.10
3	A	2889	VAL	CB-CA-C	-6.88	98.32	111.40
3	A	2458	PRO	N-CA-CB	5.79	110.25	103.30
3	A	2741	ILE	CB-CA-C	-5.75	100.10	111.60
3	A	2905	LEU	CA-CB-CG	-5.63	102.34	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	3072	TYR	Sidechain

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	C	121	0	73	10	0
2	B	335	0	262	46	0
3	A	4742	0	4783	706	0
All	All	5198	0	5118	730	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 71.

The worst 5 of 730 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)
2:B:56:LEU:HD23	2:B:57:ARG:H	1.02	1.17
3:A:2604:ALA:HB1	3:A:2677:ALA:HB3	1.27	1.17
3:A:3040:ALA:HB1	3:A:3043:LEU:HD11	1.27	1.12
3:A:2691:ALA:HB1	3:A:2692:PRO:HD2	1.25	1.11
2:B:10:LEU:HD11	3:A:2566:ARG:HH21	1.12	1.10

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	
2	B	36/70 (51%)	23 (64%)	5 (14%)	8 (22%)	0	1
3	A	594/649 (92%)	372 (63%)	117 (20%)	105 (18%)	0	2
All	All	630/719 (88%)	395 (63%)	122 (19%)	113 (18%)	0	2

5 of 113 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	17	ASP
2	B	23	PRO
2	B	38	VAL
2	B	41	ASP
2	B	43	TRP

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	
2	B	33/63 (52%)	27 (82%)	6 (18%)	2	11
3	A	521/572 (91%)	435 (84%)	86 (16%)	2	15
All	All	554/635 (87%)	462 (83%)	92 (17%)	2	15

5 of 92 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	2731	SER
3	A	2749	LEU
3	A	3078	ASN
3	A	2732	ASP
3	A	2741	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	2945	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	2956	GLN
3	A	3060	HIS
3	A	2779	GLN
3	A	3042	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 carbohydrates 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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.