wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:53 PM GMT

PDB ID	:	1MIU
Title	:	Structure of a BRCA2-DSS1 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	:	2002-08-23
Resolution	:	3.10 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 (i) symbol.

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

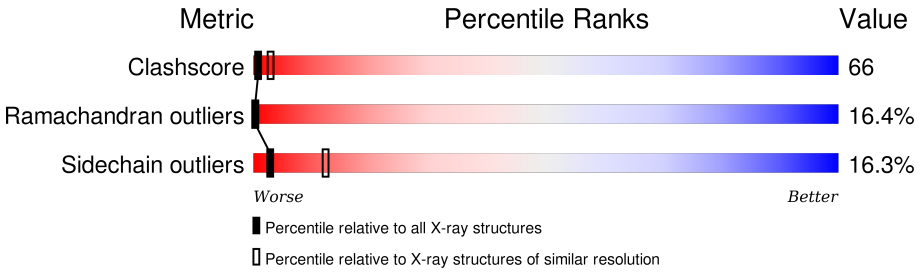
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.10 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	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)

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 <=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	B	70	23% 29% 6% • 40%
2	A	738	23% 47% 19% • 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5647 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 Deleted in split hand/split foot protein 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	B	42	348	217	53	78	0	0	0

- Molecule 2 is a protein called Breast Cancer type 2 susceptibility protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	671	5294	3361	925	991	17	0	0	0

- Molecule 3 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Hg		
3	A	5	5	5	0	0

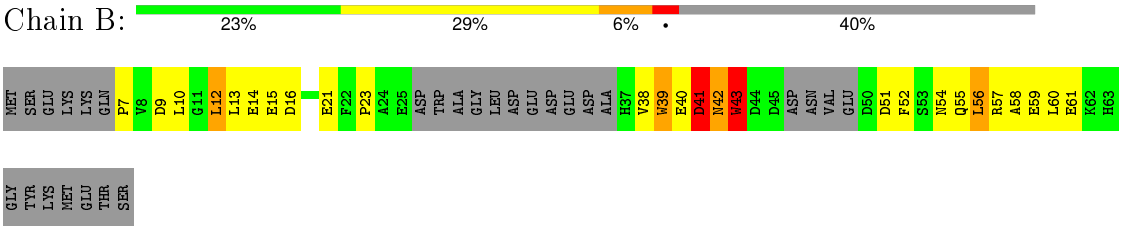
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

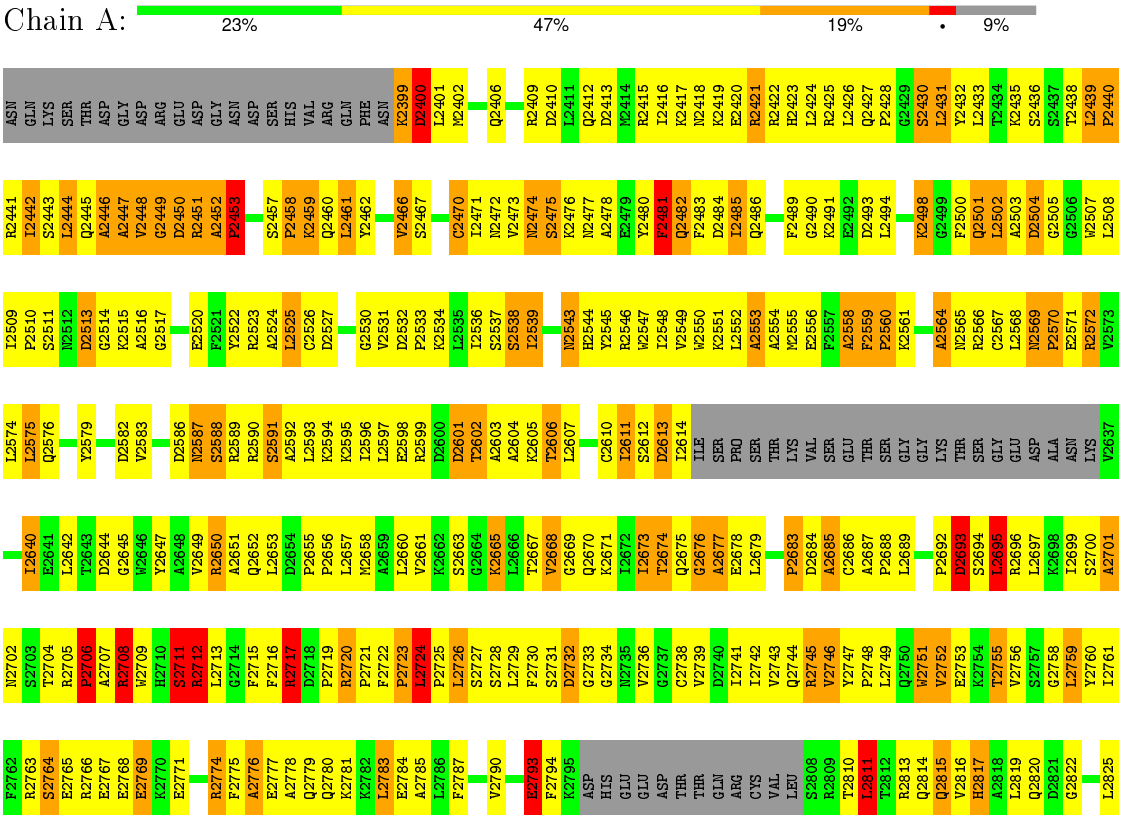
- Molecule 1: Deleted in split hand/split foot protein 1

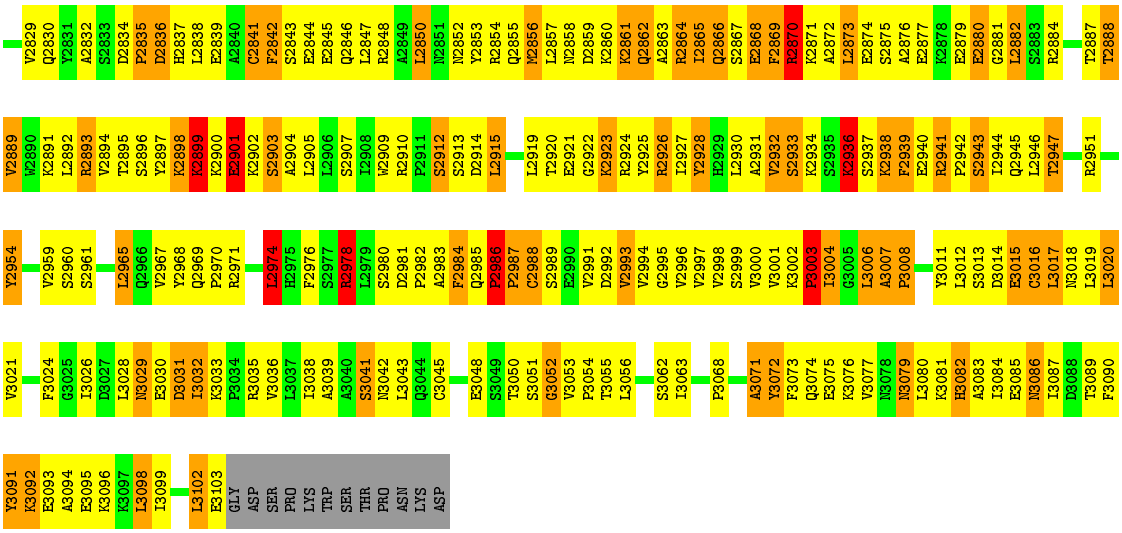
Chain B:



- Molecule 2: Breast Cancer type 2 susceptibility protein

Chain A:




V2829 Q2830 Y2831 A2832 S2833 D2834 P2835 D2836 H2837 L2838 E2839 A2840 C2841 F2842 S2843 E2844 E2845 Q2846 L2847 R2848 A2849 L2850 N2851 N2852 Y2853 R2854 Q2855 M2856 L2857 N2858 D2859 K2860 K2861 Q2862 A2863 R2864 I2865 Q2866 S2867 E2868 F2869 R2870 K2871 A2872 L2873 E2874 S2875 A2876 E2877 K2878 E2879 E2880 G2881 L2882 S2883 R2884 T2887 T2888
V2889 W2890 K2891 L2892 R2893 V2894 T2895 S2896 Y2897 K2898 K2899 K2900 E2901 K2902 S2903 A2904 L2905 L2906 S2907 I2908 W2909 R2910 P2911 S2912 S2913 D2914 L2915 L2919 T2920 E2921 G2922 K2923 R2924 Y2925 R2926 I2927 Y2928 H2929 L2930 A2931 V2932 S2933 K2934 S2935 K2936 S2937 K2938 F2939 E2940 R2941 P2942 S2943 I2944 Q2945 L2946 T2947 R2951
Y2954 V2959 S2960 S2961 L2965 Q2966 V2967 Y2968 Q2969 P2970 R2971 L2974 H2975 F2976 S2977 R2978 L2979 S2980 D2981 P2982 A2983 F2984 Q2985 P2986 P2987 C2988 S2989 E2990 V2991 D2992 V2993 V2994 G2995 V2996 V2997 V2998 S2999 V3000 V3001 K3002 P3003 I3004 G3005 L3006 A3007 P3008 Y3011 L3012 S3013 D3014 E3015 C3016 L3017 N3018 L3019 L3020
V3021 F3024 G3025 I3026 D3027 L3028 N3029 E3030 D3031 I3032 K3033 P3034 R3035 V3036 L3037 I3038 A3039 A3040 S3041 N3042 L3043 Q3044 C3045 E3048 S3049 T3050 S3051 G3052 V3053 P3054 T3055 L3056 S3062 I3063 P3068 A3071 Y3072 F3073 Q3074 E3075 K3076 V3077 N3078 N3079 L3080 K3081 H3082 A3083 I3084 E3085 N3086 I3087 D3088 T3089 F3090
Y3091 K3092 E3093 A3094 E3095 K3096 K3097 L3098 I3099 L3102 E3103 GLY ASP SER PRO LYS TRP SER THR PRO ASN LYS ASP

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	160.51Å 228.27Å 81.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.10	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-3.10)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.256 , 0.310	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5647	wwPDB-VP
Average B, all atoms (Å^2)	107.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	B	0.57	0/354	0.85	2/478 (0.4%)
2	A	0.59	6/5399 (0.1%)	0.86	9/7303 (0.1%)
All	All	0.59	6/5753 (0.1%)	0.85	11/7781 (0.1%)

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
2	A	0	1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2712	ARG	CZ-NH1	6.55	1.41	1.33
2	A	2712	ARG	CZ-NH2	6.13	1.41	1.33
2	A	2978	ARG	CZ-NH1	5.45	1.40	1.33
2	A	2978	ARG	CZ-NH2	5.41	1.40	1.33
2	A	3045	CYS	CB-SG	5.15	1.91	1.82

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	2978	ARG	NE-CZ-NH1	11.80	126.20	120.30
2	A	2458	PRO	N-CA-CB	5.95	110.44	103.30
2	A	2811	LEU	CA-CB-CG	5.82	128.69	115.30
2	A	2724	LEU	CA-CB-CG	5.76	128.54	115.30
1	B	7	PRO	N-CA-CB	5.45	109.83	103.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	2853	TYR	Sidechain

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	B	348	0	281	59	0
2	A	5294	0	5280	722	0
3	A	5	0	0	0	0
All	All	5647	0	5561	744	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 66.

The worst 5 of 744 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:A:2936:LYS:HD3	2:A:2936:LYS:H	1.02	1.15
2:A:3001:VAL:HG12	2:A:3003:PRO:HD3	1.34	1.05
2:A:2426:LEU:H	2:A:2426:LEU:HD12	1.20	1.01
1:B:9:ASP:HB3	1:B:12:LEU:HD12	1.42	0.99
2:A:3020:LEU:HA	2:A:3054:PRO:HD2	1.47	0.97

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	B	36/70 (51%)	25 (69%)	7 (19%)	4 (11%)	0 3
2	A	665/738 (90%)	427 (64%)	127 (19%)	111 (17%)	0 0
All	All	701/808 (87%)	452 (64%)	134 (19%)	115 (16%)	0 0

5 of 115 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	41	ASP
1	B	56	LEU
2	A	2440	PRO
2	A	2451	ARG
2	A	2453	PRO

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	B	36/63 (57%)	32 (89%)	4 (11%)	8 29
2	A	571/649 (88%)	476 (83%)	95 (17%)	3 12
All	All	607/712 (85%)	508 (84%)	99 (16%)	3 12

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	2746	VAL
2	A	2841	CYS
2	A	3041	SER
2	A	2752	VAL
2	A	2783	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	2814	GLN
2	A	2855	GLN
2	A	3044	GLN
2	A	2670	GLN
2	A	2675	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 5 ligands modelled in this entry, 5 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

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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.