wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 11:47 AM EDT

PDB ID	:	2B5M
Title	:	Crystal Structure of DDB1
Authors	:	Li, T.; Chen, X.; Garbutt, K.C.; Zhou, P.; Zheng, N.
Deposited on	:	2005-09-28
Resolution	:	2.92 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 (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:

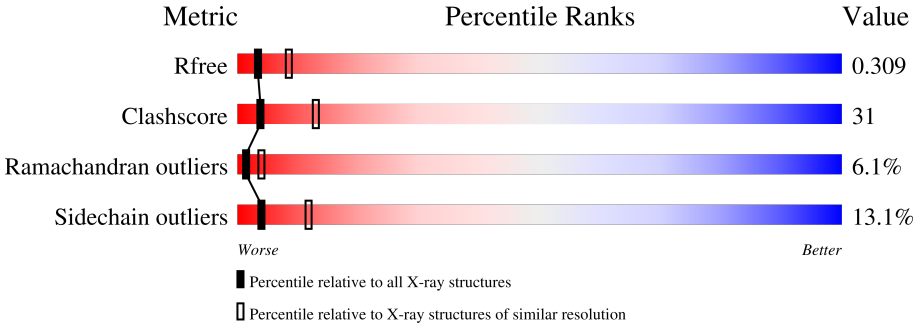
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.92 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}	130704	2307 (2.94-2.90)
Clashscore	141614	2531 (2.94-2.90)
Ramachandran outliers	138981	2462 (2.94-2.90)
Sidechain outliers	138945	2464 (2.94-2.90)

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%

Mol	Chain	Length	Quality of chain
1	A	1140	42% 40% 14% • •

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 8768 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 damage-specific DNA binding protein 1.

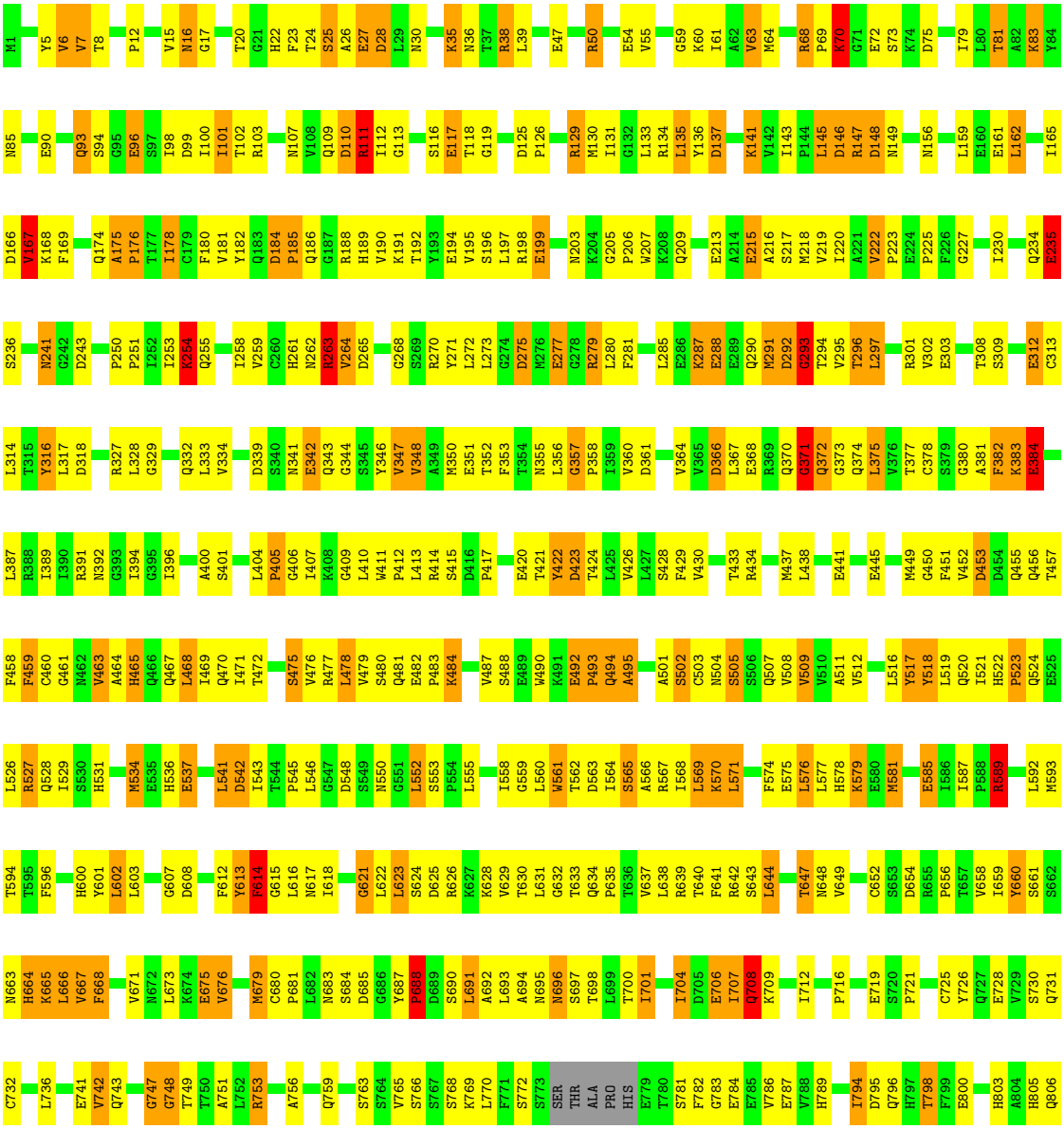
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1119	8768	5556	1477	1687	48	0	0	0

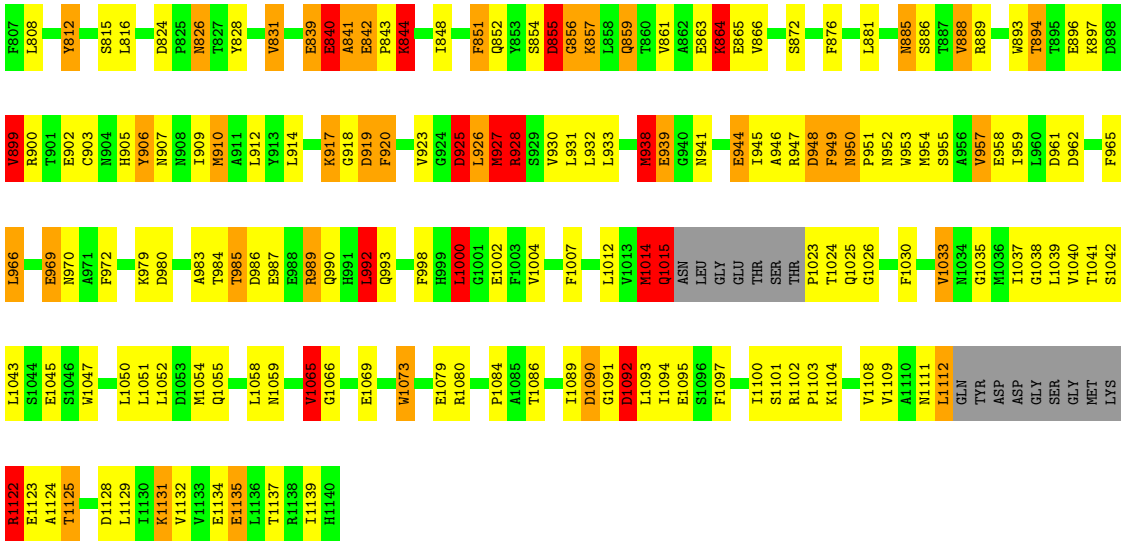
3 Residue-property plots (i)

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. 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. 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: damage-specific DNA binding protein 1

Chain A:




F807 L808 Y812 S815 L816 D824 P825 N826 T827 Y828 V831 E839 E840 A841 E842 P843 K844 I848 F851 Q852 Y853 S854 D855 G856 K857 L858 Q859 T860 V861 A862 E863 K864 E865 V866 S872 F876 L881 N885 S886 T887 V888 R889 W893 T894 T895 E896 K897 D898
V899 R900 T901 E902 C903 N904 H905 Y906 N907 N908 I909 M910 A911 L912 Y913 L914 K917 G918 D919 F920 V923 G924 D925 L926 M927 R928 S929 V930 L931 L932 L933 M938 E939 G940 N941 E944 I945 A946 R947 D948 F949 N950 P951 N952 W953 M954 S955 A956 V957 E958 I959 L960 D961 D962 F965
L966 E969 N970 A971 F972 K979 D980 A983 T984 T985 D986 E987 E988 R989 Q990 H991 L992 Q993 F998 H999 L1000 G1001 E1002 F1003 V1004 F1007 L1012 V1013 M1014 Q1015 ASN LEU GLY GLU THR SER THR P1023 T1024 Q1025 G1026 F1030 V1033 M1034 G1035 M1036 I1037 G1038 L1039 V1040 T1041 S1042
L1043 S1044 E1045 S1046 W1047 L1050 L1051 L1052 D1053 M1054 Q1055 L1058 N1059 V1065 G1066 E1069 W1073 E1079 R1080 P1084 A1085 T1086 I1089 D1090 G1091 D1092 L1093 I1094 E1095 S1096 F1097 I1100 S1101 R1102 P1103 K1104 V1108 V1109 A1110 N1111 L1112 GLN TYR ASP ASP GLY SER GLY MET LYS
R1122 E1123 A1124 T1125 D1128 L1129 I1130 K1131 V1132 V1133 E1134 E1135 L1136 T1137 R1138 I1139 H1140

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.38Å 133.83Å 184.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.70 – 2.92 48.68 – 2.92	Depositor EDS
% Data completeness (in resolution range)	(Not available) (48.70-2.92) 92.0 (48.68-2.92)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 2.91Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.232 , 0.277 0.293 , 0.309	Depositor DCC
R_{free} test set	1591 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å^2)	57.4	Xtriage
Anisotropy	0.123	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 21.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	8768	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% 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 (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	1.62	110/8928 (1.2%)	1.41	85/12091 (0.7%)

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
1	A	1	2

The worst 5 of 110 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1033	VAL	CB-CG1	-10.01	1.31	1.52
1	A	277	GLU	CD-OE1	9.55	1.36	1.25
1	A	1135	GLU	CD-OE2	8.49	1.34	1.25
1	A	841	ALA	CA-CB	8.46	1.70	1.52
1	A	857	LYS	CD-CE	8.37	1.72	1.51

The worst 5 of 85 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	986	ASP	CB-CA-C	21.05	152.49	110.40
1	A	987	GLU	N-CA-CB	-18.34	77.60	110.60
1	A	842	GLU	N-CA-CB	-15.89	82.00	110.60
1	A	841	ALA	CB-CA-C	15.19	132.88	110.10
1	A	309	SER	N-CA-CB	-12.03	92.46	110.50

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	841	ALA	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1023	PRO	Peptide
1	A	292	ASP	Peptide

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	8768	0	8744	542	0
All	All	8768	0	8744	542	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 31.

The worst 5 of 542 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:1054:MET:SD	1:A:1054:MET:CE	2.05	1.45
1:A:581:MET:SD	1:A:581:MET:CE	2.07	1.43
1:A:927:MET:SD	1:A:927:MET:CE	2.13	1.37
1:A:781:SER:HB3	1:A:784:GLU:OE1	1.31	1.27
1:A:81:THR:HG21	1:A:85:ASN:HD22	1.07	1.09

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	1111/1140 (98%)	888 (80%)	155 (14%)	68 (6%)	1 4

5 of 68 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	235	GLU
1	A	291	MET
1	A	405	PRO
1	A	417	PRO
1	A	430	VAL

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	982/999 (98%)	853 (87%)	129 (13%)	4 12

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1014	MET
1	A	1052	LEU
1	A	437	MET
1	A	434	ARG
1	A	1086	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	797	HIS
1	A	941	ASN
1	A	1070	HIS
1	A	343	GLN
1	A	332	GLN

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 monosaccharides 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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.