Full wwPDB X-ray Structure Validation Report (i)

May 26, 2020 – 10:39 am BST

PDB ID	:	2BMY
Title	:	Banana Lectin
Authors	:	Meagher, J.L.; Winter, H.C.; Ezell, P.; Goldstein, I.J.; Stuckey, J.A.
Deposited on	:	2005-03-17
Resolution	:	2.50 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 following versions of software and data (see references (i)) were used in the production of this report:

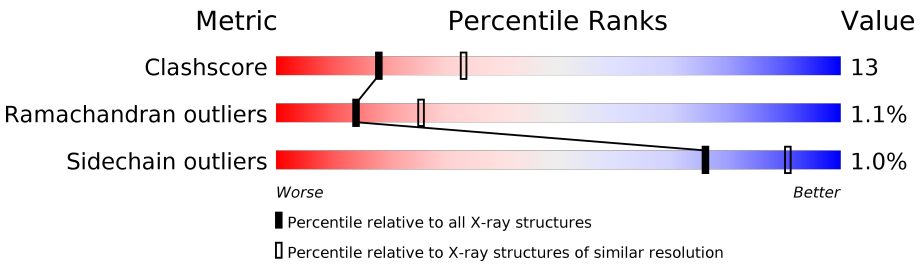
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	141	72% 26% ••
1	B	141	75% 22% •••

2 Entry composition (i)

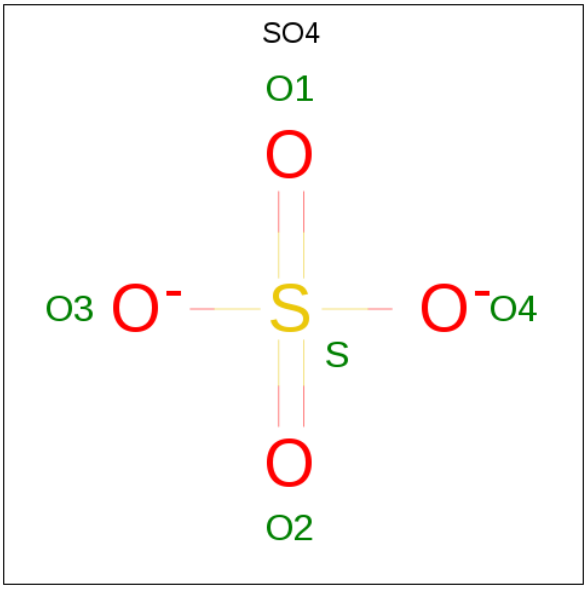
There are 4 unique types of molecules in this entry. The entry contains 2188 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 RIPENING-ASSOCIATED PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	141	1031	664	171	193	3	0	0	0
1	B	139	1014	654	168	190	2	0	0	0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cd		
3	B	1	1	1	0	0
			Total	Cd		
3	A	2	2	2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	60	60	60	0	0
			Total	O		
4	B	70	70	70	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. 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.

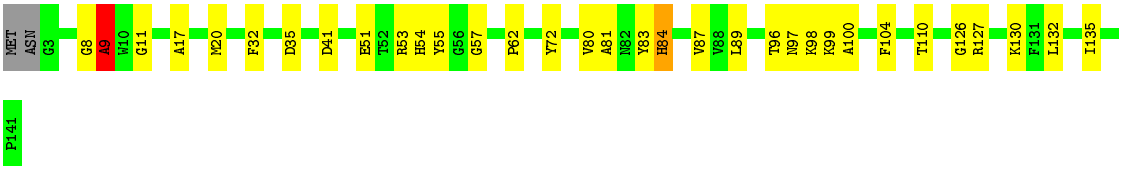
Note EDS was not executed.

- Molecule 1: RIPENING-ASSOCIATED PROTEIN

Chain A:

- Molecule 1: RIPENING-ASSOCIATED PROTEIN

Chain B:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	81.50Å 81.50Å 147.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	9.99 – 2.50	Depositor
% Data completeness (in resolution range)	98.3 (9.99-2.50)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.248 , 0.283	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2188	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

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

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	0.68	2/1058 (0.2%)	0.69	1/1427 (0.1%)
1	B	0.45	0/1040	0.71	1/1403 (0.1%)
All	All	0.58	2/2098 (0.1%)	0.70	2/2830 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	100	ALA	C-N	15.77	1.70	1.34
1	A	103	PRO	C-N	-7.73	1.16	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	9	ALA	N-CA-C	-6.32	93.92	111.00
1	B	9	ALA	N-CA-C	-5.43	96.34	111.00

There are no chirality outliers.

There are no planarity outliers.

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	1031	0	1000	30	0
1	B	1014	0	980	25	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	5	0	0	0	0
2	B	5	0	0	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
4	A	60	0	0	6	0
4	B	70	0	0	0	0
All	All	2188	0	1980	54	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 13.

All (54) 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:100:ALA:C	1:A:101:TYR:N	1.70	1.42
1:B:104:PHE:CE1	1:B:104:PHE:CE2	2.40	1.03
1:B:72:TYR:H	1:B:97:ASN:HD21	1.31	0.76
1:B:72:TYR:H	1:B:97:ASN:ND2	1.84	0.76
1:A:65:ILE:HG12	1:A:101:TYR:CD2	2.25	0.72
1:A:54:HIS:HB3	1:B:84:HIS:CE1	2.26	0.71
1:B:51:GLU:HB3	1:B:53:ARG:NH2	2.05	0.71
1:B:81:ALA:O	1:B:87:VAL:HA	1.91	0.71
1:B:32:PHE:CE2	1:B:62:PRO:HG3	2.26	0.69
1:B:51:GLU:HB3	1:B:53:ARG:HH21	1.64	0.63
4:A:2027:HOH:O	1:B:84:HIS:CD2	2.52	0.61
1:B:89:LEU:HD11	1:B:135:ILE:HG22	1.85	0.59
1:B:57:GLY:HA3	1:B:130:LYS:O	2.03	0.59
1:A:31:ILE:N	1:A:31:ILE:HD12	2.22	0.55
1:A:89:LEU:HD21	1:A:135:ILE:HG22	1.87	0.55
1:A:72:TYR:H	1:A:97:ASN:ND2	2.04	0.55
1:B:11:GLY:HA2	1:B:80:VAL:HG13	1.88	0.55
1:A:1:MET:CE	1:A:1:MET:HA	2.38	0.54
4:A:2027:HOH:O	1:B:84:HIS:HA	2.07	0.54
1:A:22:PRO:HD3	1:A:138:TYR:CE2	2.43	0.53
1:B:8:GLY:O	1:B:9:ALA:CB	2.55	0.53
1:A:56:GLY:C	4:A:2027:HOH:O	2.47	0.52
1:A:32:PHE:HA	1:A:61:THR:O	2.09	0.51
1:A:24:TYR:O	1:A:25:ARG:HG2	2.10	0.50
1:A:96:THR:C	1:A:98:LYS:H	2.15	0.50
1:A:63:HIS:HE1	4:A:2031:HOH:O	1.96	0.49
1:B:89:LEU:HD12	1:B:132:LEU:HG	1.95	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:20:MET:HB2	1:B:55:TYR:OH	2.13	0.49
1:A:8:GLY:O	1:A:9:ALA:CB	2.60	0.48
1:B:99:LYS:HG2	1:B:100:ALA:N	2.29	0.48
1:A:1:MET:HG3	1:A:4:ALA:HB3	1.95	0.48
1:A:65:ILE:HG12	1:A:101:TYR:CG	2.50	0.47
1:A:100:ALA:C	1:A:101:TYR:CA	2.75	0.46
1:B:17:ALA:HA	1:B:126:GLY:O	2.16	0.46
1:A:81:ALA:HB2	1:A:109:GLY:HA3	1.98	0.46
1:A:68:GLN:NE2	1:A:69:GLU:H	2.14	0.45
1:A:25:ARG:HH11	1:A:25:ARG:HG2	1.82	0.45
1:B:17:ALA:HB2	1:B:127:ARG:NH1	2.33	0.44
1:A:102:GLY:HA2	1:A:104:PHE:CE2	2.52	0.44
1:B:41:ASP:OD1	1:B:54:HIS:HA	2.18	0.44
1:B:96:THR:C	1:B:98:LYS:H	2.21	0.44
1:A:49:LYS:N	1:A:49:LYS:HD3	2.32	0.44
1:A:20:MET:CE	1:A:23:ALA:HB2	2.49	0.43
1:A:99:LYS:HG2	1:A:100:ALA:N	2.34	0.42
1:A:42:VAL:O	1:A:52:THR:HA	2.20	0.42
1:A:100:ALA:CA	1:A:101:TYR:N	2.70	0.41
1:A:72:TYR:H	1:A:97:ASN:HD21	1.67	0.41
1:A:53:ARG:HB3	1:A:55:TYR:CZ	2.55	0.41
1:B:35:ASP:OD2	1:B:83:TYR:OH	2.28	0.41
1:B:99:LYS:CG	1:B:100:ALA:N	2.84	0.41
1:B:32:PHE:CD2	1:B:62:PRO:HG3	2.56	0.41
1:A:63:HIS:CE1	4:A:2031:HOH:O	2.71	0.40
1:B:72:TYR:N	1:B:97:ASN:HD21	2.08	0.40
1:A:80:VAL:HG21	4:A:2046:HOH:O	2.22	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	139/141 (99%)	130 (94%)	8 (6%)	1 (1%)	22	39
1	B	137/141 (97%)	129 (94%)	6 (4%)	2 (2%)	10	18
All	All	276/282 (98%)	259 (94%)	14 (5%)	3 (1%)	14	26

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	ALA
1	B	9	ALA
1	B	84	HIS

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	100/100 (100%)	99 (99%)	1 (1%)	76	90
1	B	98/100 (98%)	97 (99%)	1 (1%)	76	90
All	All	198/200 (99%)	196 (99%)	2 (1%)	76	90

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	B	110	THR

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

Mol	Chain	Res	Type
1	A	2	ASN
1	A	63	HIS
1	A	68	GLN
1	A	97	ASN
1	B	84	HIS
1	B	97	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)

Of 5 ligands modelled in this entry, 3 are 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	$\#\|Z\| > 2$
2	SO4	A	1142	-	4,4,4	0.26	0	6,6,6	0.11	0
2	SO4	B	1142	-	4,4,4	0.27	0	6,6,6	0.07	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 (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	100:ALA	C	101:TYR	N	1.70
1	A	103:PRO	C	104:PHE	N	1.16

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.