



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 10:00 PM GMT

PDB ID : 3CSG  
Title : Crystal Structure of Monobody YS1(MBP-74)/Maltose Binding Protein Fusion Complex  
Authors : Gilbreth, R.N.; Koide, S.  
Deposited on : 2008-04-09  
Resolution : 1.80 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

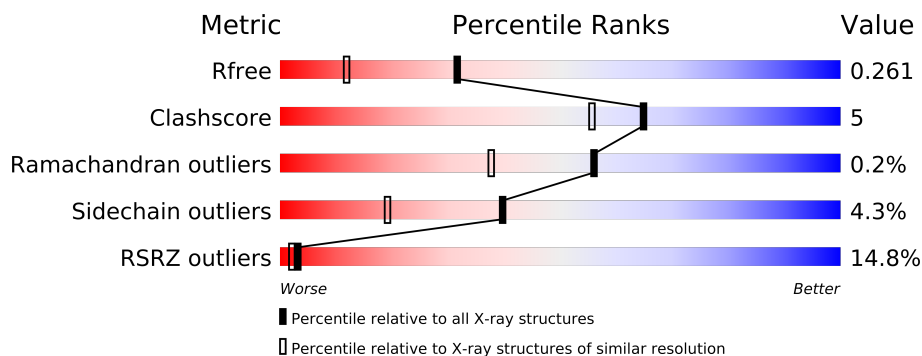
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.80 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))
$R_{free}$	66092	3513 (1.80-1.80)
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)
RSRZ outliers	66119	3515 (1.80-1.80)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	461	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3838 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Maltose-binding protein Monobody YS1 Fusion.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	458	Total	C	N	O	S	0	0	0
			3554	2290	568	690	6			

- Molecule 2 is water.

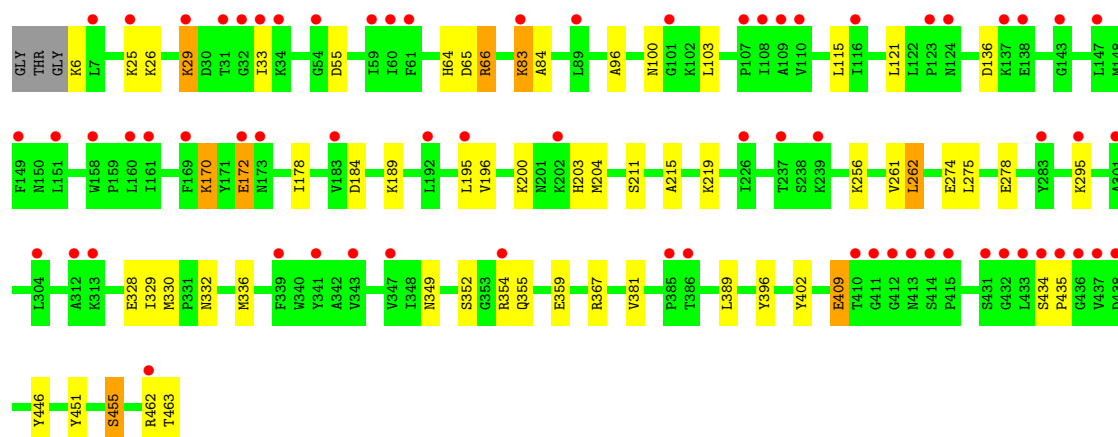
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	284	Total	O	0	0
			284	284		

### 3 Residue-property plots

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.

- Molecule 1: Maltose-binding protein Monobody YS1 Fusion

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.58Å 68.58Å 108.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.80 19.73 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (20.00-1.80) 99.6 (19.73-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 1.80Å)	Xtriage
Refinement program	CNS/REFMAC 5.2	Depositor
R, $R_{free}$	0.187 , 0.235 0.225 , 0.261	Depositor DCC
$R_{free}$ test set	4613 reflections (10.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.3	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 46.6	EDS
Estimated twinning fraction	0.049 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 46046 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3838	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

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	0.99	4/3646 (0.1%)	0.90	4/4966 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	402	TYR	CG-CD2	5.33	1.46	1.39
1	A	446	TYR	CE1-CZ	-5.32	1.31	1.38
1	A	451	TYR	CD2-CE2	5.29	1.47	1.39
1	A	359	GLU	CG-CD	5.06	1.59	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	172	GLU	C-N-CA	-5.68	107.49	121.70
1	A	262	LEU	CB-CG-CD1	5.49	120.34	111.00
1	A	336	MET	CG-SD-CE	5.45	108.92	100.20
1	A	455	SER	CB-CA-C	5.13	119.84	110.10

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3554	0	3490	37	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	284	0	0	9	0
All	All	3838	0	3490	37	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (37) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:6:LYS:HB2	1:A:33:ILE:HG23	1.46	0.98
1:A:66:ARG:HD2	2:A:1143:HOH:O	1.72	0.90
1:A:96:ALA:HB2	1:A:329:ILE:HD12	1.54	0.89
1:A:64:HIS:HD2	1:A:261:VAL:H	1.26	0.84
1:A:64:HIS:CD2	1:A:261:VAL:H	2.08	0.70
1:A:136:ASP:OD2	1:A:203:HIS:HD2	1.83	0.62
1:A:381:VAL:HG22	1:A:389:LEU:HB3	1.82	0.60
1:A:328:GLU:HB2	2:A:1203:HOH:O	2.02	0.59
1:A:100:ASN:ND2	2:A:1151:HOH:O	2.34	0.58
1:A:64:HIS:HE1	1:A:330:MET:O	1.89	0.56
1:A:256:LYS:HE2	1:A:328:GLU:HG2	1.88	0.55
1:A:204:MET:CE	2:A:1020:HOH:O	2.54	0.54
1:A:178:ILE:HD11	1:A:396:TYR:CE1	2.42	0.54
1:A:349:ASN:HD22	1:A:355:GLN:NE2	2.05	0.53
1:A:274:GLU:O	1:A:278:GLU:HG2	2.08	0.53
1:A:215:ALA:HB1	1:A:219:LYS:HE2	1.93	0.51
1:A:55:ASP:HB3	2:A:1241:HOH:O	2.11	0.49
1:A:170:LYS:HD3	1:A:170:LYS:HA	1.71	0.49
1:A:204:MET:HE2	2:A:1020:HOH:O	2.12	0.48
1:A:381:VAL:HG21	1:A:389:LEU:HD22	1.96	0.48
1:A:196:VAL:HG12	1:A:200:LYS:HD2	1.96	0.48
1:A:6:LYS:HG3	1:A:33:ILE:HD12	1.95	0.48
1:A:349:ASN:HD22	1:A:355:GLN:HE21	1.62	0.47
1:A:409:GLU:CD	1:A:409:GLU:H	2.19	0.46
1:A:435:PRO:HA	1:A:463:THR:HG22	1.98	0.46
1:A:26:LYS:O	1:A:29:LYS:HB3	2.17	0.45
1:A:178:ILE:HD11	1:A:396:TYR:CZ	2.52	0.44
1:A:352:SER:OG	1:A:354:ARG:HG3	2.18	0.43
1:A:367:ARG:NH1	2:A:1183:HOH:O	2.51	0.43
1:A:65:ASP:OD2	1:A:66:ARG:N	2.52	0.43
1:A:100:ASN:HA	2:A:1151:HOH:O	2.19	0.42
1:A:83:LYS:HE3	1:A:84:ALA:HB2	2.00	0.42
1:A:184:ASP:O	1:A:189:LYS:HE2	2.19	0.42
1:A:434:SER:HA	1:A:435:PRO:HD2	1.84	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:275:LEU:HA	1:A:275:LEU:HD23	1.80	0.42
1:A:29:LYS:HE3	2:A:1277:HOH:O	2.19	0.41
1:A:6:LYS:CB	1:A:33:ILE:HG23	2.34	0.40

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	A	456/461 (99%)	441 (97%)	14 (3%)	1 (0%)	56 38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	455	SER

### 5.3.2 Protein sidechains ⓘ

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	374/375 (100%)	358 (96%)	16 (4%)	40 19

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LYS
1	A	29	LYS
1	A	66	ARG
1	A	83	LYS

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	103	LEU
1	A	115	LEU
1	A	121	LEU
1	A	170	LYS
1	A	172	GLU
1	A	195	LEU
1	A	211	SER
1	A	262	LEU
1	A	295	LYS
1	A	332	ASN
1	A	409	GLU
1	A	462	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	ASN
1	A	49	GLN
1	A	64	HIS
1	A	86	GLN
1	A	100	ASN
1	A	201	ASN
1	A	203	HIS
1	A	218	ASN
1	A	332	ASN
1	A	355	GLN
1	A	413	ASN
1	A	460	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	458/461 (99%)	0.93	68 (14%) <b>3</b> <b>2</b>	20, 37, 47, 56	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	385	PRO	4.6
1	A	414	SER	4.5
1	A	437	VAL	4.3
1	A	413	ASN	4.1
1	A	411	GLY	4.0
1	A	415	PRO	4.0
1	A	83	LYS	3.7
1	A	160	LEU	3.6
1	A	25	LYS	3.6
1	A	183	VAL	3.6
1	A	462	ARG	3.5
1	A	33	ILE	3.4
1	A	161	ILE	3.4
1	A	354	ARG	3.3
1	A	123	PRO	3.3
1	A	60	ILE	3.3
1	A	438	ASP	3.2
1	A	343	VAL	3.2
1	A	339	PHE	3.1
1	A	29	LYS	3.1
1	A	116	ILE	3.1
1	A	436	GLY	3.0
1	A	34	LYS	3.0
1	A	433	LEU	3.0
1	A	295	LYS	3.0
1	A	301	ALA	2.9
1	A	435	PRO	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	158	TRP	2.8
1	A	195	LEU	2.8
1	A	239	LYS	2.8
1	A	54	GLY	2.8
1	A	61	PHE	2.8
1	A	237	THR	2.7
1	A	169	PHE	2.7
1	A	386	THR	2.7
1	A	101	GLY	2.7
1	A	143	GLY	2.7
1	A	412	GLY	2.6
1	A	431	SER	2.6
1	A	283	TYR	2.5
1	A	108	ILE	2.5
1	A	7	LEU	2.5
1	A	192	LEU	2.5
1	A	173	ASN	2.5
1	A	313	LYS	2.5
1	A	31	THR	2.5
1	A	432	GLY	2.5
1	A	341	TYR	2.4
1	A	124	ASN	2.4
1	A	59	ILE	2.4
1	A	410	THR	2.4
1	A	147	LEU	2.4
1	A	151	LEU	2.4
1	A	32	GLY	2.3
1	A	347	VAL	2.3
1	A	89	LEU	2.3
1	A	149	PHE	2.3
1	A	138	GLU	2.2
1	A	172	GLU	2.2
1	A	107	PRO	2.2
1	A	312	ALA	2.2
1	A	110	VAL	2.1
1	A	109	ALA	2.1
1	A	137	LYS	2.1
1	A	202	LYS	2.1
1	A	304	LEU	2.0
1	A	226	ILE	2.0
1	A	434	SER	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.