



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 11:05 PM GMT

PDB ID : 3AJA  
Title : Crystal Structure of MSMEG\_6394  
Authors : Vivian, J.P.; Scoble, J.; Beddoe, T.; Rossjohn, J.  
Deposited on : 2010-05-27  
Resolution : 2.90 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>

---

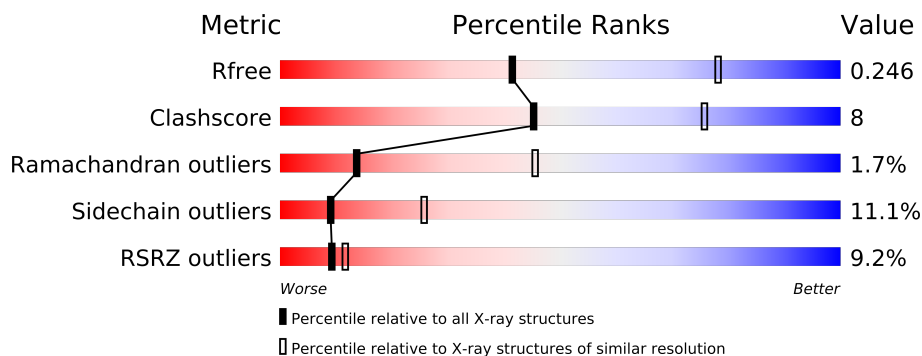
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.90 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	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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 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	302	
1	B	302	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3956 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 Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	0	0
			1952	1219	334	386	13			
1	B	264	Total	C	N	O	S	0	0	0
			1972	1230	338	391	13			

- Molecule 2 is water.

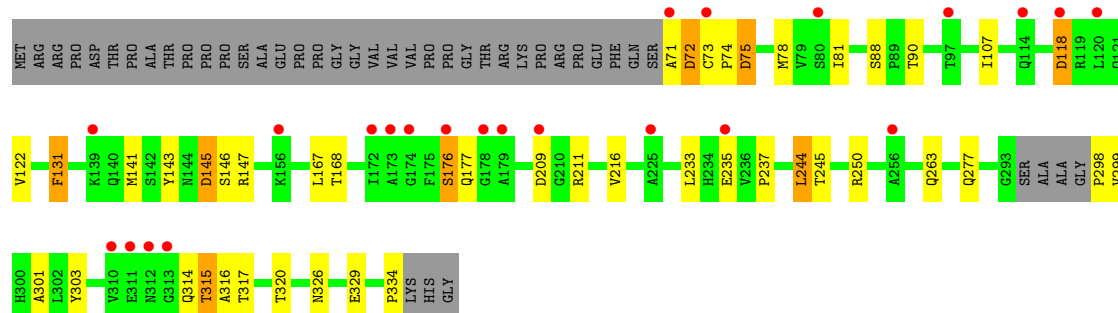
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	14	Total	O	0	0
			14	14		
2	B	18	Total	O	0	0
			18	18		

### 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.

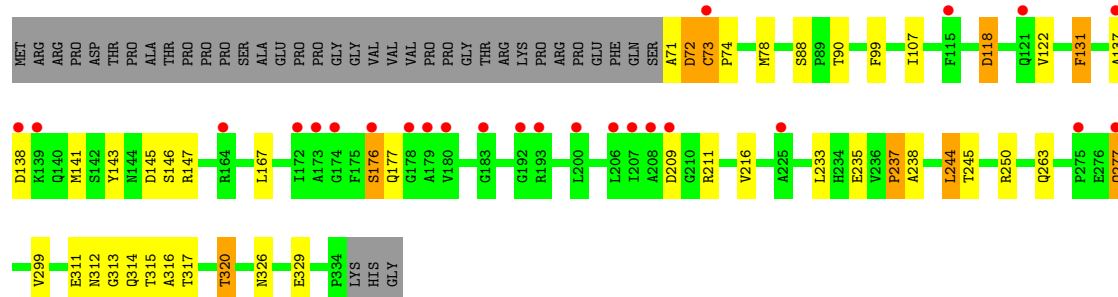
- Molecule 1: Putative uncharacterized protein

Chain A: 



- Molecule 1: Putative uncharacterized protein

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.43Å 130.43Å 209.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.17 – 2.90 29.17 – 2.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.17-2.90) 88.5 (29.17-2.90)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.80 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, $R_{free}$	0.213 , 0.250 0.210 , 0.246	Depositor DCC
$R_{free}$ test set	932 reflections (5.45%)	DCC
Wilson B-factor (Å <sup>2</sup> )	94.7	Xtriage
Anisotropy	0.018	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 79.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 18030 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3956	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	103.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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	0.63	0/2000	0.66	1/2732 (0.0%)
1	B	0.68	1/2021 (0.0%)	0.67	1/2763 (0.0%)
All	All	0.65	1/4021 (0.0%)	0.66	2/5495 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	277	GLN	CD-OE1	5.09	1.35	1.24

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	244	LEU	CA-CB-CG	6.02	129.15	115.30
1	A	244	LEU	CA-CB-CG	5.69	128.40	115.30

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	1952	0	0	16	0
1	B	1972	0	0	17	0
2	A	14	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	18	0	0	5	0
All	All	3956	0	0	33	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 8.

All (33) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:73:CYS:CB	1:A:74:PRO:CD	2.43	0.96
1:A:315:THR:CG2	1:A:316:ALA:N	2.45	0.79
1:B:72:ASP:O	2:B:25:HOH:O	2.00	0.78
1:B:71:ALA:O	1:B:72:ASP:CB	2.33	0.76
1:A:209:ASP:OD2	1:A:250:ARG:NH2	2.18	0.76
1:B:209:ASP:OD2	1:B:250:ARG:NH2	2.22	0.72
1:A:75:ASP:O	1:A:168:THR:CG2	2.39	0.71
1:B:311:GLU:O	1:B:312:ASN:C	2.30	0.70
1:B:316:ALA:O	1:B:320:THR:CG2	2.45	0.64
1:A:71:ALA:O	1:A:72:ASP:CB	2.46	0.63
1:A:145:ASP:N	1:A:145:ASP:OD1	2.30	0.62
1:A:141:MET:CE	1:A:146:SER:CA	2.83	0.56
1:B:141:MET:CE	1:B:146:SER:CA	2.85	0.55
1:A:298:PRO:O	1:A:301:ALA:N	2.40	0.55
1:A:316:ALA:O	1:A:320:THR:CG2	2.56	0.54
1:A:176:SER:OG	1:A:177:GLN:N	2.40	0.53
1:B:311:GLU:O	1:B:313:GLY:N	2.44	0.50
1:B:143:TYR:CE2	1:B:147:ARG:NH2	2.81	0.49
1:A:131:PHE:CD1	1:A:143:TYR:CD1	3.01	0.49
1:A:145:ASP:O	1:A:146:SER:C	2.54	0.46
1:B:176:SER:OG	1:B:177:GLN:N	2.50	0.45
1:B:99:PHE:N	2:B:23:HOH:O	2.50	0.45
1:B:138:ASP:CA	2:B:2:HOH:O	2.65	0.45
1:B:131:PHE:CD1	1:B:143:TYR:CD1	3.05	0.44
1:B:72:ASP:N	2:B:31:HOH:O	2.50	0.44
1:A:143:TYR:CE2	1:A:147:ARG:NH2	2.86	0.43
1:A:118:ASP:N	1:A:118:ASP:OD1	2.51	0.43
1:B:137:ALA:C	2:B:2:HOH:O	2.57	0.43
1:A:71:ALA:N	1:A:334:PRO:C	2.73	0.41
1:A:303:TYR:O	1:A:315:THR:CG2	2.69	0.41
1:B:118:ASP:N	1:B:118:ASP:OD1	2.54	0.41
1:B:237:PRO:O	1:B:238:ALA:C	2.59	0.41
1:B:73:CYS:CB	1:B:74:PRO:CD	2.98	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	256/302 (85%)	239 (93%)	12 (5%)	5 (2%)	11	40
1	B	262/302 (87%)	239 (91%)	19 (7%)	4 (2%)	15	50
All	All	518/604 (86%)	478 (92%)	31 (6%)	9 (2%)	14	45

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	ASP
1	B	72	ASP
1	A	176	SER
1	A	75	ASP
1	B	176	SER
1	A	237	PRO
1	B	237	PRO
1	B	299	VAL
1	A	299	VAL

### 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	212/245 (86%)	189 (89%)	23 (11%)	9	27
1	B	213/245 (87%)	189 (89%)	24 (11%)	9	25
All	All	425/490 (87%)	378 (89%)	47 (11%)	9	26



All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	78	MET
1	A	81	ILE
1	A	88	SER
1	A	90	THR
1	A	107	ILE
1	A	118	ASP
1	A	122	VAL
1	A	131	PHE
1	A	145	ASP
1	A	167	LEU
1	A	211	ARG
1	A	216	VAL
1	A	233	LEU
1	A	235	GLU
1	A	244	LEU
1	A	245	THR
1	A	263	GLN
1	A	277	GLN
1	A	314	GLN
1	A	315	THR
1	A	317	THR
1	A	326	ASN
1	A	329	GLU
1	B	73	CYS
1	B	78	MET
1	B	88	SER
1	B	90	THR
1	B	107	ILE
1	B	118	ASP
1	B	122	VAL
1	B	131	PHE
1	B	145	ASP
1	B	167	LEU
1	B	211	ARG
1	B	216	VAL
1	B	233	LEU
1	B	235	GLU
1	B	244	LEU
1	B	245	THR
1	B	263	GLN
1	B	277	GLN
1	B	314	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	315	THR
1	B	317	THR
1	B	320	THR
1	B	326	ASN
1	B	329	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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	260/302 (86%)	0.49	23 (8%) 10 13	82, 101, 133, 162	0
1	B	264/302 (87%)	0.53	25 (9%) 8 11	76, 97, 140, 174	0
All	All	524/604 (86%)	0.51	48 (9%) 9 11	76, 99, 138, 174	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	179	ALA	5.1
1	B	174	GLY	4.4
1	B	176	SER	4.2
1	B	173	ALA	4.1
1	B	208	ALA	3.8
1	B	137	ALA	3.8
1	B	200	LEU	3.7
1	A	172	ILE	3.7
1	A	120	LEU	3.6
1	B	121	GLN	3.3
1	B	138	ASP	3.3
1	A	173	ALA	3.3
1	A	174	GLY	3.2
1	B	206	LEU	3.0
1	A	311	GLU	3.0
1	A	156	LYS	2.9
1	A	312	ASN	2.8
1	A	235	GLU	2.8
1	B	180	VAL	2.8
1	B	164	ARG	2.7
1	B	139	LYS	2.7
1	A	178	GLY	2.7
1	B	178	GLY	2.7
1	B	192	GLY	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	97	THR	2.5
1	B	183	GLY	2.5
1	A	209	ASP	2.4
1	A	313	GLY	2.4
1	B	209	ASP	2.4
1	A	73	CYS	2.4
1	A	80	SER	2.4
1	A	176	SER	2.4
1	B	275	PRO	2.4
1	A	139	LYS	2.3
1	A	118	ASP	2.3
1	B	277	GLN	2.3
1	A	179	ALA	2.3
1	A	310	VAL	2.3
1	B	193	ARG	2.3
1	B	207	ILE	2.3
1	B	73	CYS	2.2
1	A	256	ALA	2.2
1	B	225	ALA	2.1
1	A	114	GLN	2.1
1	A	225	ALA	2.1
1	B	172	ILE	2.0
1	B	115	PHE	2.0
1	A	71	ALA	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.