



wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 02:48 PM GMT

PDB ID : 4M1A
Title : Crystal structure of a Domain of unknown function (DUF1904) from *Sebadella termitidis* ATCC 33386
Authors : Kumaran, D.; Chamala, S.; Evans, B.; Foti, R.; Gizzi, A.; Hillerich, B.; Kar, A.; Lafleur, J.; Seidel, R.; Villigas, G.; Zencheck, W.; Al Obaidi, N.; Almo, S.C.; Swaminathan, S.; New York Structural Genomics Research Consortium (NYSGRC)
Deposited on : 2013-08-02
Resolution : 1.90 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

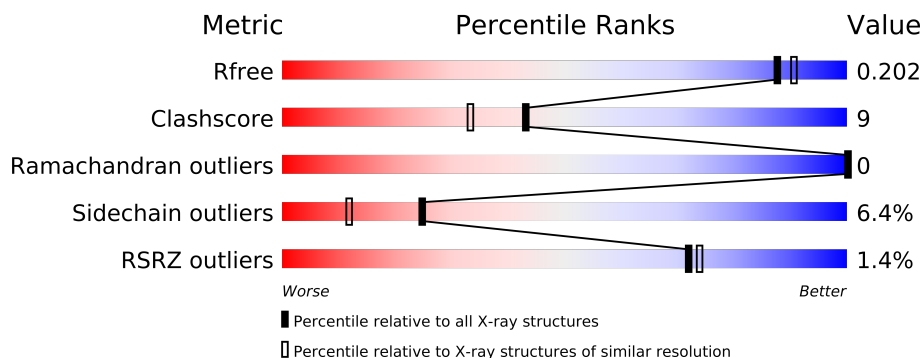
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.90 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	3684 (1.90-1.90)
Clashscore	79885	4465 (1.90-1.90)
Ramachandran outliers	78287	4413 (1.90-1.90)
Sidechain outliers	78261	4414 (1.90-1.90)
RSRZ outliers	66119	3686 (1.90-1.90)

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 ≥ 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	126	
1	B	126	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1893 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 Hypothetical protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	108	Total	C	N	O	S	Se	0	0	0
			898	586	142	167	1	2			
1	B	108	Total	C	N	O	S	Se	0	0	0
			898	586	142	167	1	2			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	109	ALA	-	EXPRESSION TAG	UNP D1AJN4
A	110	GLU	-	EXPRESSION TAG	UNP D1AJN4
A	111	ASN	-	EXPRESSION TAG	UNP D1AJN4
A	112	LEU	-	EXPRESSION TAG	UNP D1AJN4
A	113	TYR	-	EXPRESSION TAG	UNP D1AJN4
A	114	PHE	-	EXPRESSION TAG	UNP D1AJN4
A	115	GLN	-	EXPRESSION TAG	UNP D1AJN4
A	116	GLY	-	EXPRESSION TAG	UNP D1AJN4
A	117	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	118	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	119	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	120	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	121	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	122	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	123	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	124	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	125	HIS	-	EXPRESSION TAG	UNP D1AJN4
A	126	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	109	ALA	-	EXPRESSION TAG	UNP D1AJN4
B	110	GLU	-	EXPRESSION TAG	UNP D1AJN4
B	111	ASN	-	EXPRESSION TAG	UNP D1AJN4
B	112	LEU	-	EXPRESSION TAG	UNP D1AJN4
B	113	TYR	-	EXPRESSION TAG	UNP D1AJN4
B	114	PHE	-	EXPRESSION TAG	UNP D1AJN4
B	115	GLN	-	EXPRESSION TAG	UNP D1AJN4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	116	GLY	-	EXPRESSION TAG	UNP D1AJN4
B	117	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	118	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	119	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	120	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	121	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	122	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	123	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	124	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	125	HIS	-	EXPRESSION TAG	UNP D1AJN4
B	126	HIS	-	EXPRESSION TAG	UNP D1AJN4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	51	Total O 51 51	0	0
2	B	46	Total O 46 46	0	0

i

- Molecule 1: Hypothetical protein

[illegible]

- Molecule 1: Hypothetical protein

Rank	Country	Score
1	USA	100.0
2	USA	99.0
3	USA	98.0
4	USA	97.0
5	USA	96.0
6	USA	95.0
7	USA	94.0
8	USA	93.0
9	USA	92.0
10	USA	91.0
11	USA	90.0
12	USA	89.0
13	USA	88.0
14	USA	87.0
15	USA	86.0
16	USA	85.0
17	USA	84.0
18	USA	83.0
19	USA	82.0
20	USA	81.0
21	USA	80.0
22	USA	79.0
23	USA	78.0
24	USA	77.0
25	USA	76.0
26	USA	75.0
27	USA	74.0
28	USA	73.0
29	USA	72.0
30	USA	71.0
31	USA	70.0
32	USA	69.0
33	USA	68.0
34	USA	67.0
35	USA	66.0
36	USA	65.0
37	USA	64.0
38	USA	63.0
39	USA	62.0
40	USA	61.0
41	USA	60.0
42	USA	59.0
43	USA	58.0
44	USA	57.0
45	USA	56.0
46	USA	55.0
47	USA	54.0
48	USA	53.0
49	USA	52.0
50	USA	51.0
51	USA	50.0
52	USA	49.0
53	USA	48.0
54	USA	47.0
55	USA	46.0
56	USA	45.0
57	USA	44.0
58	USA	43.0
59	USA	42.0
60	USA	41.0
61	USA	40.0
62	USA	39.0
63	USA	38.0
64	USA	37.0
65	USA	36.0
66	USA	35.0
67	USA	34.0
68	USA	33.0
69	USA	32.0
70	USA	31.0
71	USA	30.0
72	USA	29.0
73	USA	28.0
74	USA	27.0
75	USA	26.0
76	USA	25.0
77	USA	24.0
78	USA	23.0
79	USA	22.0
80	USA	21.0
81	USA	20.0
82	USA	19.0
83	USA	18.0
84	USA	17.0
85	USA	16.0
86	USA	15.0
87	USA	14.0
88	USA	13.0
89	USA	12.0
90	USA	11.0
91	USA	10.0
92	USA	9.0
93	USA	8.0
94	USA	7.0
95	USA	6.0
96	USA	5.0
97	USA	4.0
98	USA	3.0
99	USA	2.0
100	USA	1.0
101	USA	0.0
102	USA	0.0
103	USA	0.0
104	USA	0.0
105	USA	0.0
106	USA	0.0
107	USA	0.0
108	USA	0.0
109	USA	0.0
110	USA	0.0
111	USA	0.0
112	USA	0.0
113	USA	0.0
114	USA	0.0
115	USA	0.0
116	USA	0.0
117	USA	0.0
118	USA	0.0
119	USA	0.0
120	USA	0.0
121	USA	0.0
122	USA	0.0
123	USA	0.0
124	USA	0.0
125	USA	0.0
126	USA	0.0
127	USA	0.0
128	USA	0.0
129	USA	0.0
130	USA	0.0
131	USA	0.0
132	USA	0.0
133	USA	0.0
134	USA	0.0
135	USA	0.0
136	USA	0.0
137	USA	0.0
138	USA	0.0
139	USA	0.0
140	USA	0.0
141	USA	0.0
142	USA	0.0
143	USA	0.0
144	USA	0.0
145	USA	0.0
146	USA	0.0
147	USA	0.0
148	USA	0.0
149	USA	0.0
150	USA	0.0
151	USA	

4 Data and refinement statistics

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α , β , γ	90.28Å 90.28Å 37.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.18 – 1.90 45.14 – 1.90	Depositor EDS
% Data completeness (in resolution range)	94.7 (45.18-1.90) 94.7 (45.14-1.90)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	8.38 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.176 , 0.206 0.183 , 0.202	Depositor DCC
R_{free} test set	1291 reflections (5.39%)	DCC
Wilson B-factor (Å ²)	23.6	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 27.5	EDS
Estimated twinning fraction	0.489 for -h,-k,l 0.026 for h,-h-k,-l 0.026 for -k,-h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 25228 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	1893	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.79% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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	1.12	0/921	1.19	9/1244 (0.7%)
1	B	1.13	0/921	1.22	10/1244 (0.8%)
All	All	1.12	0/1842	1.20	19/2488 (0.8%)

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	70	ARG	NE-CZ-NH2	-10.51	115.05	120.30
1	B	70	ARG	NE-CZ-NH1	9.60	125.10	120.30
1	A	70	ARG	NE-CZ-NH1	8.19	124.40	120.30
1	A	70	ARG	NE-CZ-NH2	-7.36	116.62	120.30
1	B	65	ARG	NE-CZ-NH1	7.30	123.95	120.30

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	898	0	862	17	0
1	B	898	0	862	17	0
2	A	51	0	0	0	0
2	B	46	0	0	1	0
All	All	1893	0	1724	31	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 9.

The worst 5 of 31 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1:MSE:SE	1:A:65:ARG:HH12	1.90	1.04
1:B:1:MSE:SE	1:B:65:ARG:HH12	2.10	0.85
1:A:1:MSE:HG2	1:A:2:PRO:N	2.07	0.68
1:A:1:MSE:SE	1:A:65:ARG:NH1	2.73	0.67
1:A:1:MSE:HG3	1:A:62:TRP:HA	1.77	0.67

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	106/126 (84%)	106 (100%)	0	0	100	100
1	B	106/126 (84%)	106 (100%)	0	0	100	100
All	All	212/252 (84%)	212 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

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	94/108 (87%)	89 (95%)	5 (5%)	32	18
1	B	94/108 (87%)	87 (93%)	7 (7%)	20	8
All	All	188/216 (87%)	176 (94%)	12 (6%)	25	12

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1	MSE
1	B	4	ILE
1	B	55	LEU
1	A	62	TRP
1	B	53	ASP

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, 95th 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 ²)	Q<0.9
1	A	108/126 (85%)	-0.25	1 (0%) 81 83	16, 25, 46, 63	0
1	B	108/126 (85%)	-0.24	2 (1%) 64 65	15, 25, 46, 60	0
All	All	216/252 (85%)	-0.25	3 (1%) 72 74	15, 25, 46, 63	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MSE	5.8
1	B	1	MSE	4.9
1	B	54	GLY	2.1

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.