wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 02:04 am BST

PDB ID	:	1ZQ7
Title	:	X-Ray Crystal Structure of Protein Q8PZK8 from Methanosarcina mazei. Northeast Structural Genomics Consortium Target MaR9.
Authors	:	Kuzin, A.P.; Abashidze, M.; Vorobiev, S.M.; Yong, W.; Xiao, R.; Ma, L.-C.; Acton, T.; Montelione, G.T.; Tong, L.; Hunt, J.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2005-05-18
Resolution	:	2.11 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 following versions of software and data (see references (1)) were used in the production of this report:

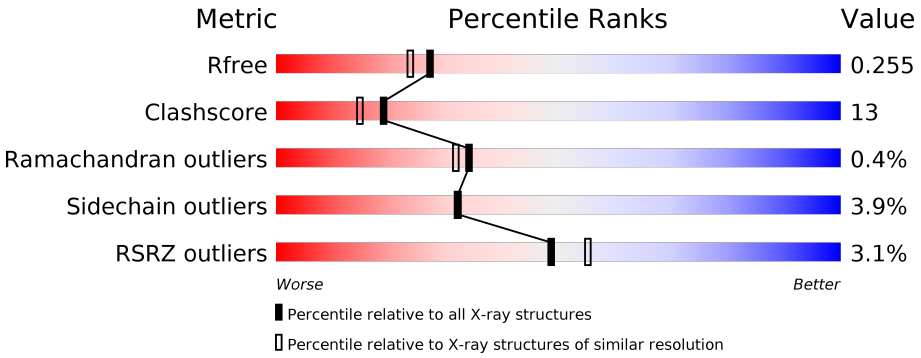
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.11 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	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	207	5% 71% 22%
1	B	207	% 77% 19%
1	C	207	4% 69% 27%
1	D	207	% 69% 29%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6583 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 Hypothetical protein MM0484.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	201	Total	C	N	O	S	Se	0	0	0
			1564	994	260	302	4	4			
1	B	204	Total	C	N	O	S	Se	0	0	0
			1584	1006	263	307	4	4			
1	C	204	Total	C	N	O	S	Se	0	0	0
			1584	1006	263	307	4	4			
1	D	204	Total	C	N	O	S	Se	0	0	0
			1584	1006	263	307	4	4			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
A	96	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
A	151	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
A	162	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
A	200	LEU	-	CLONING ARTIFACT	UNP Q8PZK8
A	201	GLU	-	CLONING ARTIFACT	UNP Q8PZK8
A	202	HIS	-	EXPRESSION TAG	UNP Q8PZK8
A	203	HIS	-	EXPRESSION TAG	UNP Q8PZK8
A	204	HIS	-	EXPRESSION TAG	UNP Q8PZK8
A	205	HIS	-	EXPRESSION TAG	UNP Q8PZK8
A	206	HIS	-	EXPRESSION TAG	UNP Q8PZK8
A	207	HIS	-	EXPRESSION TAG	UNP Q8PZK8
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
B	96	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
B	151	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
B	162	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
B	200	LEU	-	CLONING ARTIFACT	UNP Q8PZK8
B	201	GLU	-	CLONING ARTIFACT	UNP Q8PZK8
B	202	HIS	-	EXPRESSION TAG	UNP Q8PZK8
B	203	HIS	-	EXPRESSION TAG	UNP Q8PZK8
B	204	HIS	-	EXPRESSION TAG	UNP Q8PZK8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	205	HIS	-	EXPRESSION TAG	UNP Q8PZK8
B	206	HIS	-	EXPRESSION TAG	UNP Q8PZK8
B	207	HIS	-	EXPRESSION TAG	UNP Q8PZK8
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
C	96	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
C	151	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
C	162	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
C	200	LEU	-	CLONING ARTIFACT	UNP Q8PZK8
C	201	GLU	-	CLONING ARTIFACT	UNP Q8PZK8
C	202	HIS	-	EXPRESSION TAG	UNP Q8PZK8
C	203	HIS	-	EXPRESSION TAG	UNP Q8PZK8
C	204	HIS	-	EXPRESSION TAG	UNP Q8PZK8
C	205	HIS	-	EXPRESSION TAG	UNP Q8PZK8
C	206	HIS	-	EXPRESSION TAG	UNP Q8PZK8
C	207	HIS	-	EXPRESSION TAG	UNP Q8PZK8
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
D	96	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
D	151	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
D	162	MSE	MET	MODIFIED RESIDUE	UNP Q8PZK8
D	200	LEU	-	CLONING ARTIFACT	UNP Q8PZK8
D	201	GLU	-	CLONING ARTIFACT	UNP Q8PZK8
D	202	HIS	-	EXPRESSION TAG	UNP Q8PZK8
D	203	HIS	-	EXPRESSION TAG	UNP Q8PZK8
D	204	HIS	-	EXPRESSION TAG	UNP Q8PZK8
D	205	HIS	-	EXPRESSION TAG	UNP Q8PZK8
D	206	HIS	-	EXPRESSION TAG	UNP Q8PZK8
D	207	HIS	-	EXPRESSION TAG	UNP Q8PZK8

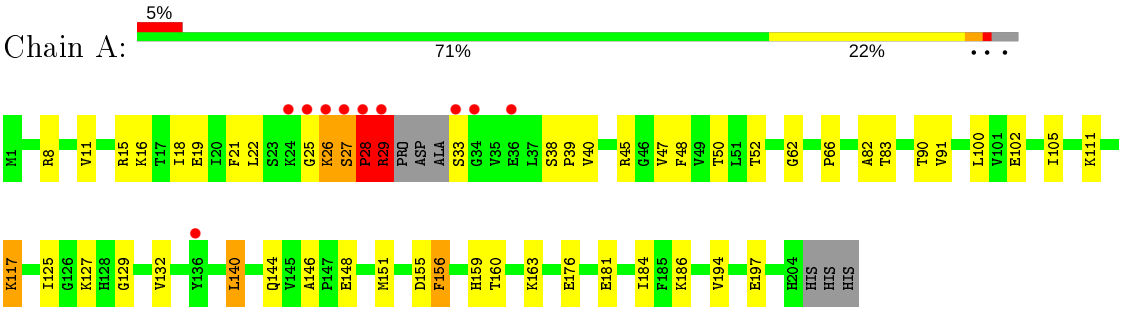
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	64	Total O 64 64	0	0
2	B	71	Total O 71 71	0	0
2	C	68	Total O 68 68	0	0
2	D	64	Total O 64 64	0	0

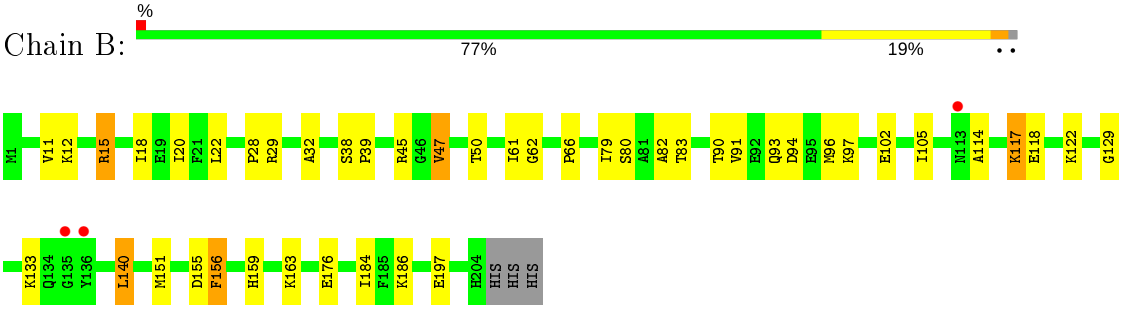
3 Residue-property plots (i)

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 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.

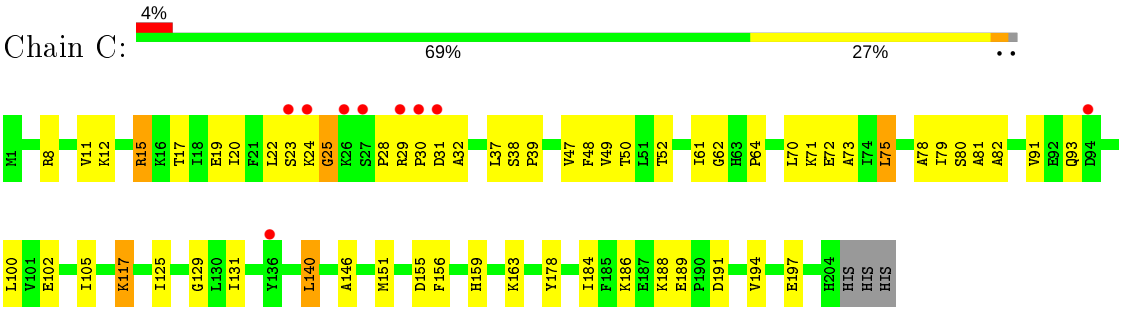
- Molecule 1: Hypothetical protein MM0484



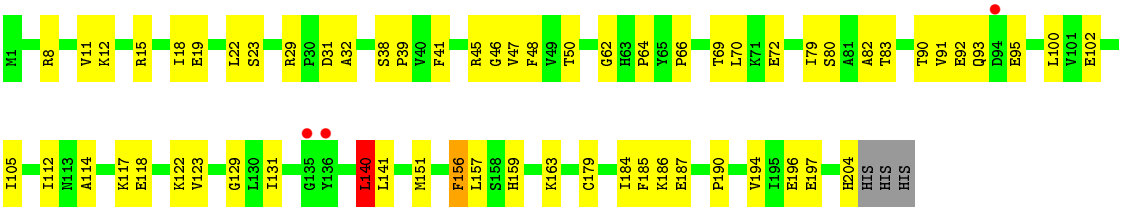
- Molecule 1: Hypothetical protein MM0484



- Molecule 1: Hypothetical protein MM0484



- Molecule 1: Hypothetical protein MM0484


M1
R8
V11
K12
R15
I18
E19
L22
S23
R29
P30
D31
A32
S38
P39
V40
F41
R45
G46
V47
F48
V49
T50
G62
H63
P64
Y65
P66
T69
L70
K71
E72
I79
S80
A81
A82
T83
T90
V91
E92
Q93
D94
E95
L100
V101
E102
I105
I112
N113
A114
K117
E118
K122
V123
G129
L130
I131
G135
Y136
L140
L141
M151
F156
L157
S158
H159
K163
C179
I184
F185
K186
E187
P190
V194
I195
E196
E197
H204
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	137.88Å 57.58Å 115.65Å 90.00° 110.60° 90.00°	Depositor
Resolution (Å)	29.83 – 2.11 29.83 – 2.12	Depositor EDS
% Data completeness (in resolution range)	76.2 (29.83-2.11) 81.9 (29.83-2.12)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.70 (at 2.12Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.208 , 0.248 0.218 , 0.255	Depositor DCC
R_{free} test set	3711 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	22.0	Xtriage
Anisotropy	0.468	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 47.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6583	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.82% 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

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	0.44	2/1592 (0.1%)	0.76	5/2145 (0.2%)
1	B	0.33	0/1614	0.62	1/2178 (0.0%)
1	C	0.34	0/1614	0.62	1/2178 (0.0%)
1	D	0.32	0/1614	0.60	2/2178 (0.1%)
All	All	0.36	2/6434 (0.0%)	0.65	9/8679 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	28	PRO	CA-C	-6.17	1.40	1.52
1	A	29	ARG	CZ-NH2	-5.08	1.26	1.33

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	28	PRO	C-N-CA	11.65	150.82	121.70
1	A	28	PRO	CA-C-N	-8.08	99.42	117.20
1	C	129	GLY	N-CA-C	-6.21	97.58	113.10
1	A	29	ARG	NE-CZ-NH2	-5.97	117.32	120.30
1	B	129	GLY	N-CA-C	-5.93	98.28	113.10

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	1564	0	1567	40	0
1	B	1584	0	1586	37	0
1	C	1584	0	1586	48	0
1	D	1584	0	1586	46	0
2	A	64	0	0	0	0
2	B	71	0	0	1	0
2	C	68	0	0	2	0
2	D	64	0	0	2	0
All	All	6583	0	6325	164	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.

The worst 5 of 164 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:26:LYS:HG2	1:A:27:SER:H	1.25	1.01
1:A:26:LYS:HG2	1:A:27:SER:N	1.78	0.95
1:A:28:PRO:O	1:A:29:ARG:HD3	1.68	0.91
1:D:47:VAL:HG12	1:D:105:ILE:HA	1.53	0.91
1:C:11:VAL:HG21	1:C:194:VAL:HG21	1.57	0.84

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	197/207 (95%)	194 (98%)	2 (1%)	1 (0%)	29 25
1	B	202/207 (98%)	199 (98%)	3 (2%)	0	100 100
1	C	202/207 (98%)	197 (98%)	4 (2%)	1 (0%)	29 25
1	D	202/207 (98%)	197 (98%)	4 (2%)	1 (0%)	29 25

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	803/828 (97%)	787 (98%)	13 (2%)	3 (0%)	34 32

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	PRO
1	C	25	GLY
1	D	23	SER

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	173/174 (99%)	164 (95%)	9 (5%)	23 20
1	B	175/174 (101%)	169 (97%)	6 (3%)	37 38
1	C	175/174 (101%)	168 (96%)	7 (4%)	31 31
1	D	175/174 (101%)	170 (97%)	5 (3%)	42 44
All	All	698/696 (100%)	671 (96%)	27 (4%)	32 32

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	140	LEU
1	C	15	ARG
1	D	156	PHE
1	B	156	PHE
1	A	47	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	93	GLN
1	D	108	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	159	HIS
1	A	203	HIS
1	C	98	ASN

5.3.3 RNA

There are no RNA molecules 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^{th} 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	197/207 (95%)	0.14	10 (5%) 28 33	10, 24, 46, 53	0
1	B	200/207 (96%)	-0.07	3 (1%) 73 77	10, 21, 37, 47	0
1	C	200/207 (96%)	0.03	9 (4%) 33 38	9, 20, 43, 51	0
1	D	200/207 (96%)	0.11	3 (1%) 73 77	13, 25, 41, 49	0
All	All	797/828 (96%)	0.05	25 (3%) 49 55	9, 23, 41, 53	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	27	SER	7.5
1	C	136	TYR	6.9
1	A	28	PRO	6.6
1	A	29	ARG	6.2
1	D	136	TYR	4.6

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 (i)

There are no such residues in this entry.