



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 06:46 am BST

PDB ID : 2Z8U
Title : Methanococcus jannaschii TBP
Authors : Adachi, N.; Senda, T.; Horikoshi, M.
Deposited on : 2007-09-10
Resolution : 1.90 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
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

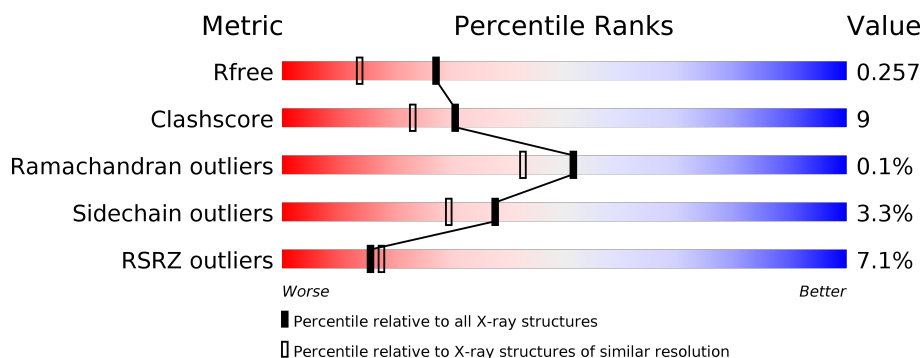
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

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}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (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 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 $\leq 5\%$. 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	188	<div> <div>16%</div> <div> <div></div> <div>63%</div> <div>25%</div> <div>• 10%</div> </div> </div>
1	B	188	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>11%</div> <div>• 7%</div> </div> </div>
1	P	188	<div> <div>5%</div> <div> <div></div> <div>79%</div> <div>12%</div> <div>• 8%</div> </div> </div>
1	Q	188	<div> <div>3%</div> <div> <div></div> <div>76%</div> <div>16%</div> <div>• 7%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5600 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 TATA-box-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	169	Total	C	N	O	S	0	0	0
			1293	826	209	253	5			
1	B	175	Total	C	N	O	S	0	0	0
			1348	864	215	264	5			
1	P	173	Total	C	N	O	S	0	0	0
			1334	854	213	262	5			
1	Q	175	Total	C	N	O	S	0	0	0
			1348	864	215	264	5			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	EXPRESSION TAG	UNP Q57930
A	-6	HIS	-	EXPRESSION TAG	UNP Q57930
A	-5	HIS	-	EXPRESSION TAG	UNP Q57930
A	-4	HIS	-	EXPRESSION TAG	UNP Q57930
A	-3	HIS	-	EXPRESSION TAG	UNP Q57930
A	-2	HIS	-	EXPRESSION TAG	UNP Q57930
A	-1	HIS	-	EXPRESSION TAG	UNP Q57930
A	0	PRO	-	EXPRESSION TAG	UNP Q57930
B	-7	MET	-	EXPRESSION TAG	UNP Q57930
B	-6	HIS	-	EXPRESSION TAG	UNP Q57930
B	-5	HIS	-	EXPRESSION TAG	UNP Q57930
B	-4	HIS	-	EXPRESSION TAG	UNP Q57930
B	-3	HIS	-	EXPRESSION TAG	UNP Q57930
B	-2	HIS	-	EXPRESSION TAG	UNP Q57930
B	-1	HIS	-	EXPRESSION TAG	UNP Q57930
B	0	PRO	-	EXPRESSION TAG	UNP Q57930
P	-7	MET	-	EXPRESSION TAG	UNP Q57930
P	-6	HIS	-	EXPRESSION TAG	UNP Q57930
P	-5	HIS	-	EXPRESSION TAG	UNP Q57930
P	-4	HIS	-	EXPRESSION TAG	UNP Q57930
P	-3	HIS	-	EXPRESSION TAG	UNP Q57930

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	-2	HIS	-	EXPRESSION TAG	UNP Q57930
P	-1	HIS	-	EXPRESSION TAG	UNP Q57930
P	0	PRO	-	EXPRESSION TAG	UNP Q57930
Q	-7	MET	-	EXPRESSION TAG	UNP Q57930
Q	-6	HIS	-	EXPRESSION TAG	UNP Q57930
Q	-5	HIS	-	EXPRESSION TAG	UNP Q57930
Q	-4	HIS	-	EXPRESSION TAG	UNP Q57930
Q	-3	HIS	-	EXPRESSION TAG	UNP Q57930
Q	-2	HIS	-	EXPRESSION TAG	UNP Q57930
Q	-1	HIS	-	EXPRESSION TAG	UNP Q57930
Q	0	PRO	-	EXPRESSION TAG	UNP Q57930

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	58	Total O 58 58	0	0
2	B	94	Total O 94 94	0	0
2	P	53	Total O 53 53	0	0
2	Q	72	Total O 72 72	0	0

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.

Chain A:

16% 63% 25% 10%

Amino Acid	Category
Met	Grey
His	Grey
His	Grey
His	Grey
His	Grey
His	Grey
Pro	Grey
Pro	Grey
Met	Grey
Met	Grey
GLU	Green
P3	Green
E4	Green
I5	Yellow
N9	Yellow
V10	Yellow
K15	Yellow
D18	Green
N19	Green
I20	Green
D21	Yellow
L22	Green
E23	Green
E24	Green
V25	Green
A26	Green
M27	Green
I28	Green
L29	Green
E30	Green
N31	Green
A32	Orange
E33	Yellow
TYR	Grey
GLU	Grey
PRO	Grey
GLU	Grey
GLU	Grey
GLN	Grey
PHE	Grey
PRO	Grey
G41	Green
L42	Green
V43	Yellow
V51	Yellow
A52	Green
L53	Green
L54	Green
I55	Green
T56	Green
F57	Green
R57	Green
S58	Green
G59	Green
K60	Green
V61	Green
K69	Yellow
I74	Green
A75	Yellow
K78	Green
I79	Green
I80	Green
L83	Green
A86	Green
G87	Green
I88	Green
D89	Green
V90	Green
I96	Green
K97	Green
I98	Green
G99	Green
N100	Green
M101	Green
V102	Green
I109	Green
M112	Orange
D115	Yellow
L118	Green
M119	Green
E124	Green
Y125	Yellow
F126	Green
P127	Yellow
E128	Green
Q129	Green
F130	Green
GLU	Green
P131	Green
V144	Green
L145	Green
L146	Green
T155	Yellow
I170	Green
L171	Green
D172	Green
T173	Green
I174	Green
K175	Green
Q178	Green
GLU	Grey
LEU	Grey

Chain B:

Category	Percentage
GLU	2%
LEU	7%
Other	81%

Chain P:

5%

79%

12%

8%

MET HIS HIS HIS HIS HIS PRO MET GLU PRO E4 I5 N9 V10 T14 E33 E37 Q38 F39 P40 V43 V48 R57 K69 E73 K77 I88 D89 V90 P94 K97 I98 Q99 N100 M101 D106 N112 D115 L118 E124 E125

Chain Q:

Category	Percentage
MET	76%
H1S	3%
H2S	16%
H3S	7%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	53.24Å 55.53Å 123.43Å 90.00° 91.05° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 61.71 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-1.90) 99.9 (61.71-1.90)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.50 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.226 , 0.259 0.225 , 0.257	Depositor DCC
R_{free} test set	2892 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	31.8	Xtriage
Anisotropy	0.428	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 47.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.012 for -k,-h,-l 0.014 for k,h,-l 0.025 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5600	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.70% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.77	4/1303 (0.3%)	0.64	0/1756
1	B	0.57	0/1363	0.63	0/1841
1	P	0.45	0/1348	0.59	0/1820
1	Q	0.51	0/1363	0.62	0/1841
All	All	0.59	4/5377 (0.1%)	0.62	0/7258

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	86	ALA	C-N	10.05	1.51	1.33
1	A	41	GLY	C-O	9.78	1.39	1.23
1	A	24	GLU	CD-OE1	7.22	1.33	1.25
1	A	24	GLU	CD-OE2	6.38	1.32	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	86	ALA	Mainchain

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	1293	0	1372	42	0
1	B	1348	0	1417	28	0
1	P	1334	0	1400	19	0
1	Q	1348	0	1417	27	0
2	A	58	0	0	6	0
2	B	94	0	0	3	0
2	P	53	0	0	2	0
2	Q	72	0	0	3	0
All	All	5600	0	5606	101	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 9.

The worst 5 of 101 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:ASN:ND2	1:B:124:GLU:HG2	1.57	1.18
1:A:100:ASN:HD21	1:B:124:GLU:CG	1.60	1.13
1:A:155:THR:HB	2:A:234:HOH:O	1.65	0.96
1:A:100:ASN:ND2	1:B:124:GLU:CG	2.23	0.94
1:B:38:GLN:HG2	2:B:270:HOH:O	1.68	0.93

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	165/188 (88%)	162 (98%)	2 (1%)	1 (1%)	25	15
1	B	173/188 (92%)	169 (98%)	4 (2%)	0	100	100
1	P	171/188 (91%)	168 (98%)	3 (2%)	0	100	100
1	Q	173/188 (92%)	172 (99%)	1 (1%)	0	100	100
All	All	682/752 (91%)	671 (98%)	10 (2%)	1 (0%)	51	43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	32	ALA

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	146/165 (88%)	138 (94%)	8 (6%)	21	12
1	B	152/165 (92%)	149 (98%)	3 (2%)	55	51
1	P	150/165 (91%)	147 (98%)	3 (2%)	55	51
1	Q	152/165 (92%)	146 (96%)	6 (4%)	32	23
All	All	600/660 (91%)	580 (97%)	20 (3%)	38	29

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	35	GLU
1	B	112	ASN
1	Q	37	GLU
1	A	173	THR
1	B	9	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	P	9	ASN
1	P	99	GLN
1	Q	99	GLN
1	B	112	ASN
1	P	129	GLN

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](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	169/188 (89%)	1.06	30 (17%) 1 1	27, 38, 48, 60	0
1	B	175/188 (93%)	0.54	3 (1%) 70 72	24, 37, 49, 52	0
1	P	173/188 (92%)	0.77	10 (5%) 23 25	33, 39, 51, 61	0
1	Q	175/188 (93%)	0.65	6 (3%) 45 48	31, 38, 50, 56	0
All	All	692/752 (92%)	0.75	49 (7%) 16 17	24, 38, 50, 61	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	29	LEU	12.5
1	A	86	ALA	6.2
1	P	39	PHE	5.8
1	A	25	VAL	5.4
1	A	57	ARG	4.9

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.