



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 11:28 am GMT

PDB ID : 2Q2E
Title : Crystal structure of the topoisomerase VI holoenzyme from *Methanosarcina mazei*
Authors : Corbett, K.D.; Benedetti, P.; Berger, J.M.
Deposited on : 2007-05-28
Resolution : 4.00 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

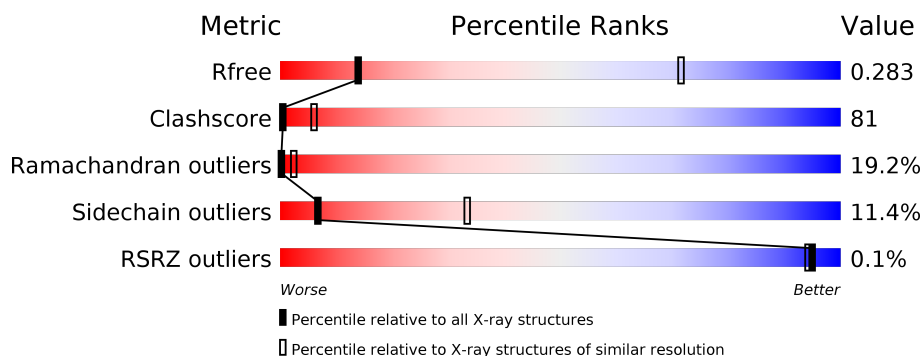
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.00 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}	100719	1088 (4.40-3.60)
Clashscore	112137	1187 (4.40-3.60)
Ramachandran outliers	110173	1139 (4.40-3.60)
Sidechain outliers	110143	1126 (4.40-3.60)
RSRZ outliers	101464	1099 (4.40-3.60)

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. 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	369	
2	B	621	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7039 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 Type II DNA topoisomerase VI subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	315	Total	C	N	O	S	0	0	0
			2532	1610	423	490	9			

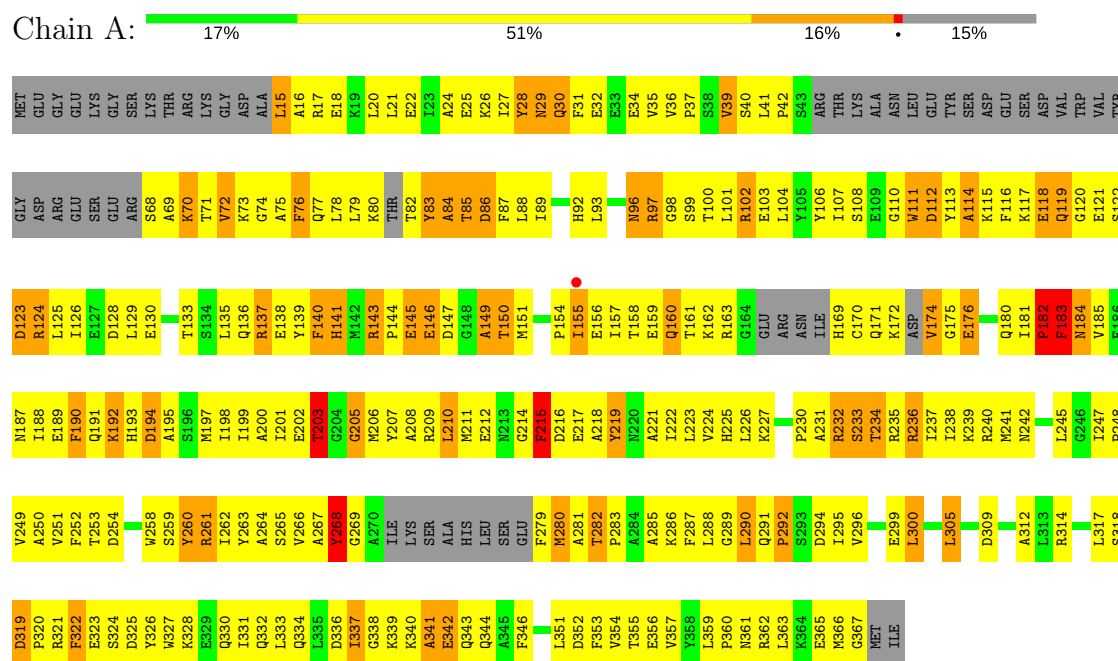
- Molecule 2 is a protein called Type 2 DNA topoisomerase 6 subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	582	Total	C	N	O	S	0	0	1
			4507	2878	768	843	18			

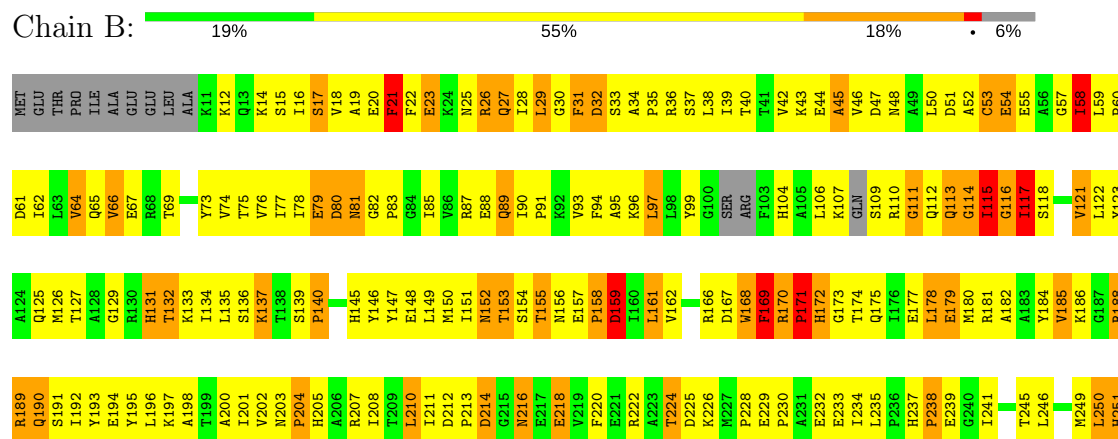
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Type II DNA topoisomerase VI subunit A



• Molecule 2: Type 2 DNA topoisomerase 6 subunit B



V561		D500	D440	V380	I320	Y252
		I501	I441	P381	G321	T253
M564		N502	P442	L382	E322	E254
		P503	V443	L383	D323	ARG
D567		V504	I444	Y384	L324	GLN
Y568		V505	K445	Q385	I325	LVS
D569		A506	E446	Q386	Y326	LEU
Y570		K507	E447	G387	R327	ALA
Y571		I508	I448	G388	G328	P260
		N509	D449	C389	L329	F261
S575		G510	L450	V390	E330	L262
A576		N511	A451	T391	K331	R263
S577		L512	I452	T392	E332	
		L513	K453	H393	T333	C267
S580		V514	E454	A394	T394	K269
S581		H515	V455	V395	V335	ILE
K582		R516	A456	E396	D336	GLY
V583		V517	R457	D397	F337	L271
L584		L518	K458	T398	I338	L272
S585			L459	K399	A339	T273
Y586		N521	K460	W400	T340	A274
K587		G522	H461	K401	S341	
I588		D523	Y462	Q402	T342	L277
E589		G524	L463	Y403	R343	C278
S590		T525	S464	G404	K344	
A591		V526	K465	L405	P345	A281
S592		D527	Q466	N406	A346	G282
E593		V528	S467	Q407	V347	L283
E594		A529	N468	P408	Y348	D284
E595		I530	L469	G409	S349	P285
L596		K531	K470	G410	G350	E286
Q597		V532	K471	G411	N351	T287
K598		K533	K472	I412	P352	D288
L599		N534	R473	P413	F353	P289
P600		PHE	E474	V414	V354	H290
Q601		GLY	K475	G415	V355	A291
L602		THR	E476	P416	E356	L292
I603		SER	I477	V417	V357	
V604		A539	I478	I418	G358	A297
E605		Y540	I479	L419	M359	
GLY		S541	T480	L420	A360	L300
ILE		F542	K481	I421	Y361	L301
GLU		R543	V482	H422	G362	E302
GLU		V544	L483	V423	G363	A303
GLU		H545	P484	A424	N364	F304
LEU		E546	K485	S425	L365	E305
VAL		M547	L486	I426	P366	K306
T613		L548	A487	N427	K367	V307
G614		P549	A488	V428	E368	K308
A615		C550	K489	P429	E369	I309
K616		K551	V490	F430	K370	K310
ALA		V552	A491	T431	I371	A311
PHE		SER	H492	S432	S372	P312
LYS		GLY	V493	E433	I373	P313
GLY		ALA	L494	S434	M374	T314
VAL		K556	E495	K435	R375	D315
		P557	K496	D436	F376	C316
		E558	D497	A437	A377	L317
		P559	V498	I438	N378	S318
		K560	P499	A439	R379	P319

4 Data and refinement statistics

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, α , β , γ	227.81Å 227.81Å 208.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 4.00 92.23 – 4.00	Depositor EDS
% Data completeness (in resolution range)	94.2 (30.00-4.00) 94.1 (92.23-4.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.20 (at 4.01Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.306 , 0.349 0.302 , 0.283	Depositor DCC
R_{free} test set	1298 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	153.1	Xtriage
Anisotropy	0.370	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 210.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	7039	wwPDB-VP
Average B, all atoms (Å ²)	192.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.40% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

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	0.41	0/2579	0.71	0/3471
2	B	0.51	1/4591 (0.0%)	0.83	5/6215 (0.1%)
All	All	0.47	1/7170 (0.0%)	0.79	5/9686 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	53	CYS	CB-SG	-6.66	1.71	1.82

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	57	GLY	N-CA-C	5.88	127.79	113.10
2	B	408	PRO	N-CA-CB	5.82	110.29	103.30
2	B	522	GLY	N-CA-C	-5.75	98.72	113.10
2	B	58	ILE	N-CA-C	5.66	126.29	111.00
2	B	362	GLY	N-CA-C	5.42	126.64	113.10

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	2532	0	2476	393	0
2	B	4507	0	4565	779	0
All	All	7039	0	7041	1140	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 81.

The worst 5 of 1140 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)
2:B:599:LEU:H	2:B:599:LEU:HD23	1.09	1.13
2:B:516:ARG:HH12	2:B:597:GLN:HB2	0.93	1.09
2:B:371:ILE:HG22	2:B:414:VAL:HG12	1.13	1.09
2:B:66:VAL:HG23	2:B:213:PRO:HD3	1.35	1.08
2:B:222:ARG:HH21	2:B:332:GLU:HB2	1.12	1.06

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	303/369 (82%)	163 (54%)	78 (26%)	62 (20%)	0	2
2	B	566/621 (91%)	339 (60%)	122 (22%)	105 (19%)	0	2
All	All	869/990 (88%)	502 (58%)	200 (23%)	167 (19%)	0	2

5 of 167 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	29	ASN
1	A	72	VAL
1	A	97	ARG
1	A	112	ASP
1	A	113	TYR

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	268/315 (85%)	240 (90%)	28 (10%)	8	37
2	B	486/527 (92%)	428 (88%)	58 (12%)	6	32
All	All	754/842 (90%)	668 (89%)	86 (11%)	7	33

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	109	SER
2	B	171	PRO
2	B	540	TYR
2	B	117	ILE
2	B	159	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	25	ASN
2	B	48	ASN
2	B	190	GLN
1	A	344	GLN
2	B	378	ASN

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	315/369 (85%)	-0.04	1 (0%) 93 91	139, 217, 258, 265	0
2	B	582/621 (93%)	-0.06	0 100 100	93, 175, 248, 265	0
All	All	897/990 (90%)	-0.05	1 (0%) 95 94	93, 192, 254, 265	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	155	ILE	2.0

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.