wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 05:15 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

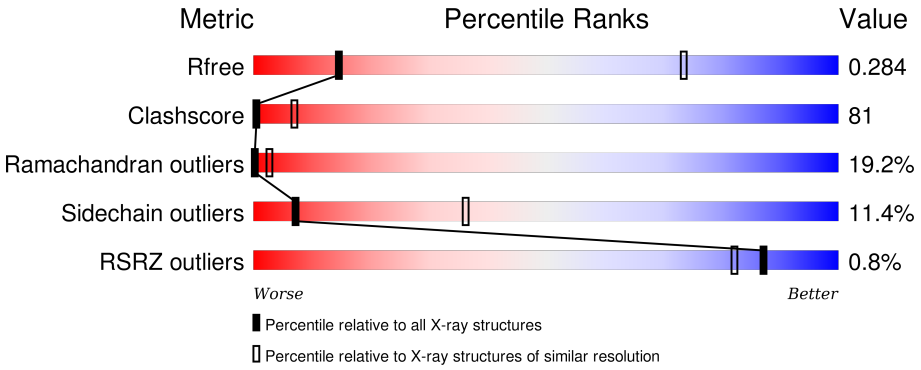
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 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}	91344	1010 (4.42-3.56)
Clashscore	102246	1052 (4.40-3.60)
Ramachandran outliers	100387	1005 (4.40-3.60)
Sidechain outliers	100360	1013 (4.42-3.58)
RSRZ outliers	91569	1013 (4.42-3.56)

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 <=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	% 17% 51% 16% • 15%
2	B	621	19% 55% 18% • 6%

2 Entry composition (i)

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
			Total	C	N	O	S			
1	A	315	2532	1610	423	490	9	0	0	0

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

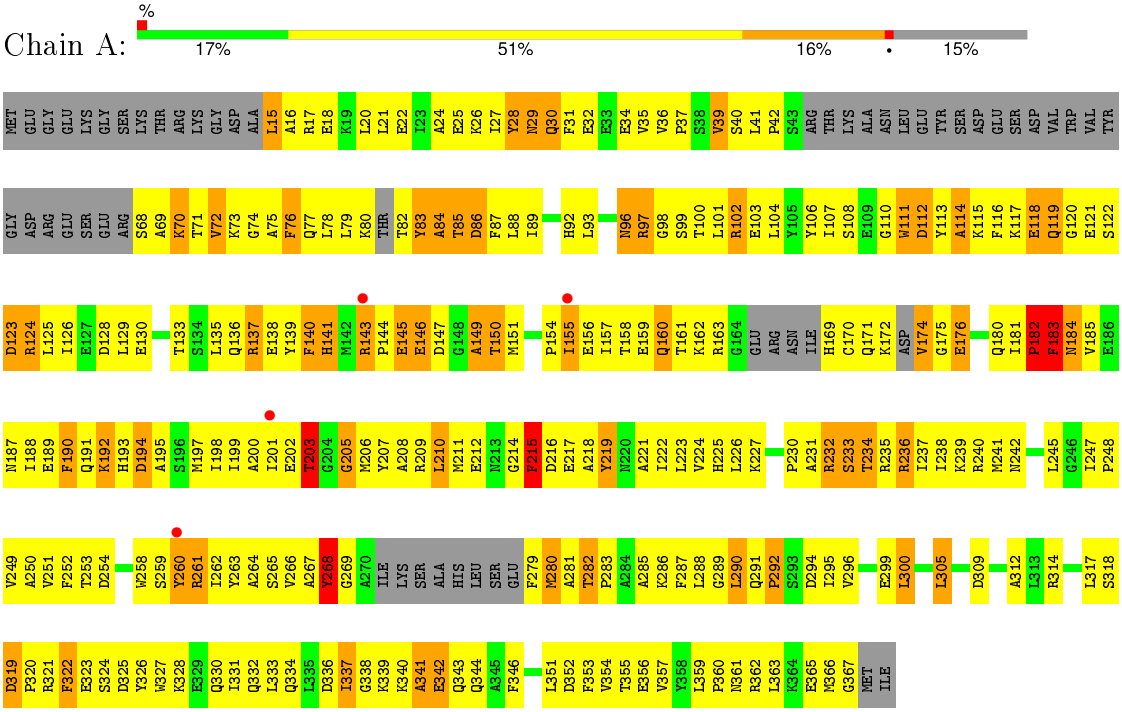
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	582	4507	2878	768	843	18	0	0	1

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 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.

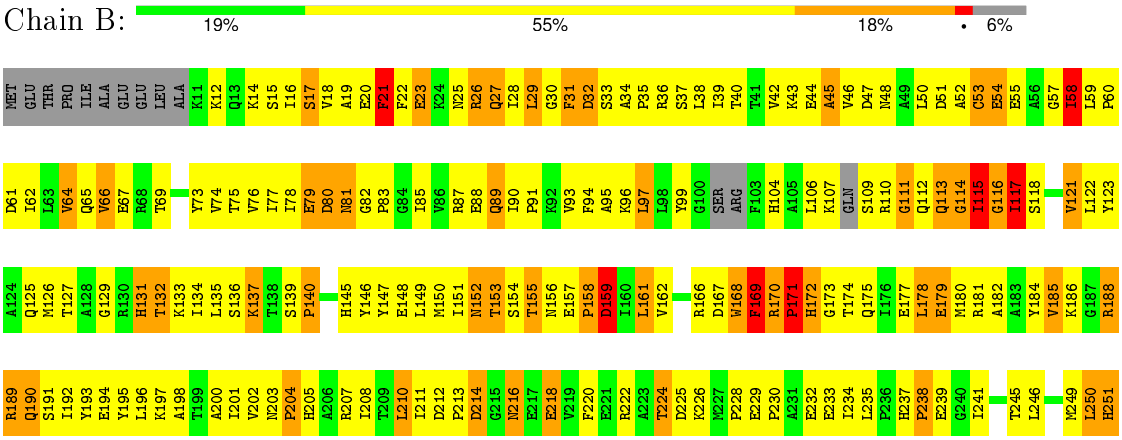
- Molecule 1: Type II DNA topoisomerase VI subunit A

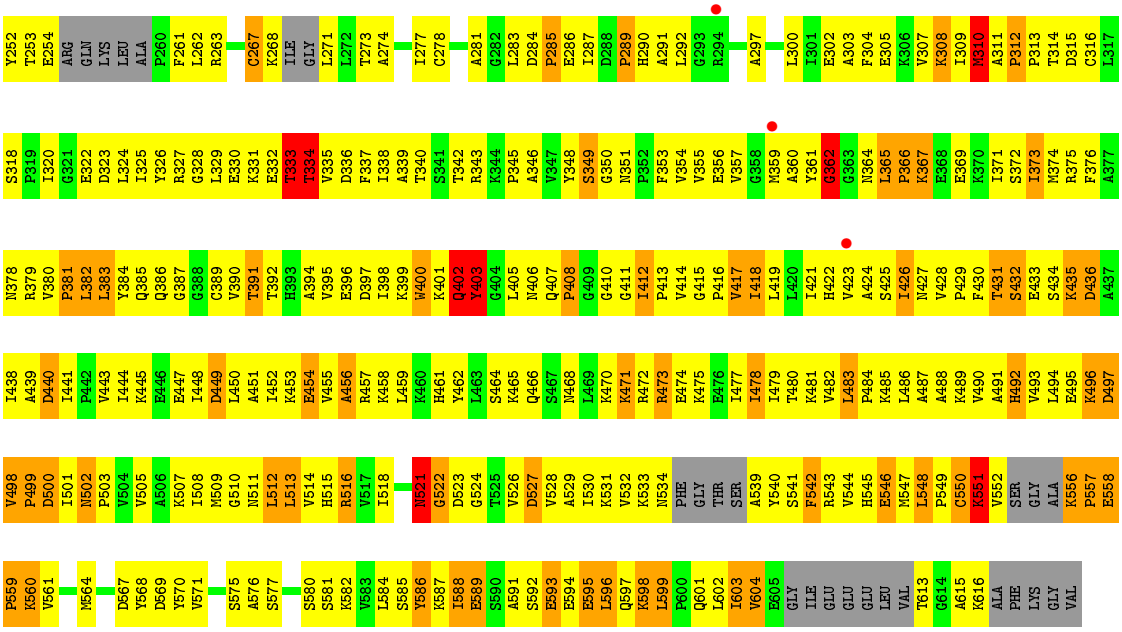
Chain A:



- Molecule 2: Type 2 DNA topoisomerase 6 subunit B

Chain B:




Y252 T253 E254 ARG GLN LYS LEU ALA P260 F261 L262 R263 C267 K268 ILE GLY L271 L272 T273 A274 I277 C278 A281 G282 L283 D284 P285 E286 I287 D288 P289 H290 A291 L292 G293 R294 A297 L300 I301 E302 A303 F304 E305 K306 V307 K308 I309 M310 A311 P312 P313 T314 D315 C316 L317
S318 P319 I320 G321 E322 D323 L324 I325 Y326 R327 G328 L329 E330 K331 E332 T333 T334 V335 D336 F337 I338 A339 T340 S341 T342 R343 K344 P345 A346 V347 Y348 S349 G350 N351 P352 F353 V354 V355 E356 V357 G358 M359 A360 Y361 G362 G363 N364 L365 P366 K367 E368 E369 K370 I371 S372 I373 M374 R375 F376 A377
N378 R379 V380 P381 L382 L383 Y384 Q385 Q386 G387 G388 C389 V390 T391 T392 H393 A394 V395 E396 D397 I398 K399 W400 K401 Q402 Y403 G404 L405 N406 Q407 P408 G409 G410 G411 I412 P413 V414 G415 P416 V417 I418 L419 L420 I421 H422 V423 A424 S425 I426 N427 V428 P429 F430 T431 S432 E433 S434 K435 D436 A437
I438 A439 D440 I441 P442 V443 I444 K445 E446 E447 I448 D449 L450 A451 I452 K453 E454 V455 A456 R457 K458 L459 K460 H461 Y462 L463 S464 K465 Q466 S467 N468 L469 K470 K471 R472 R473 E474 K475 E476 I477 I478 I479 T480 K481 V482 L483 P484 K485 L486 A487 A488 K489 V490 A491 H492 V493 L494 E495 K496 D497
V498 P499 D500 I501 N502 P503 V504 V505 A506 K507 I508 M509 G510 N511 L512 L513 V514 H515 R516 V517 I518 N521 G522 D523 G524 T525 V526 D527 V528 A529 I530 K531 V532 K533 N534 PHE GLY THR SER A539 Y540 S541 F542 R543 V544 H545 E546 M547 L548 P549 C550 K551 V552 SER GLY ALA K556 P557 E558
P559 K560 V561 M564 D567 Y568 D569 Y570 V571 S575 A576 S577 S580 S581 K582 V583 L584 S585 Y586 K587 I588 E589 S590 A591 S592 E593 E594 E595 L596 Q597 K598 L599 P600 Q601 L602 I603 V604 E605 GLY ILE GLU GLU GLU LEU VAL T613 G614 A615 K616 ALA PHE LYS GLY VAL

4 Data and refinement statistics (i)

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
$< I/\sigma(I) >$ [1]	2.20 (at 4.01Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.306 , 0.349 0.296 , 0.284	Depositor DCC
R_{free} test set	1294 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 , 211.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Outliers	0 of 25919 reflections	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.41	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(^{o})	Ideal(^{o})
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

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	3
All	All	869/990 (88%)	502 (58%)	200 (23%)	167 (19%)	0	3

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%)	9 40
2	B	486/527 (92%)	428 (88%)	58 (12%)	6 34
All	All	754/842 (90%)	668 (89%)	86 (11%)	7 36

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

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	315/369 (85%)	0.01	4 (1%)	79	70	139, 217, 258, 265	0
2	B	582/621 (93%)	-0.02	3 (0%)	91	88	93, 175, 248, 265	0
All	All	897/990 (90%)	-0.01	7 (0%)	87	82	93, 192, 254, 265	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	155	ILE	2.2
1	A	260	TYR	2.1
1	A	143	ARG	2.1
2	B	359	MET	2.1
1	A	201	ILE	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.