wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2020 – 01:51 am BST

PDB ID	:	3FPP
Title	:	Crystal structure of E.coli MacA
Authors	:	Yum, S.; Xu, Y.; Piao, S.; Ha, N.-C.
Deposited on	:	2009-01-06
Resolution	:	2.99 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:

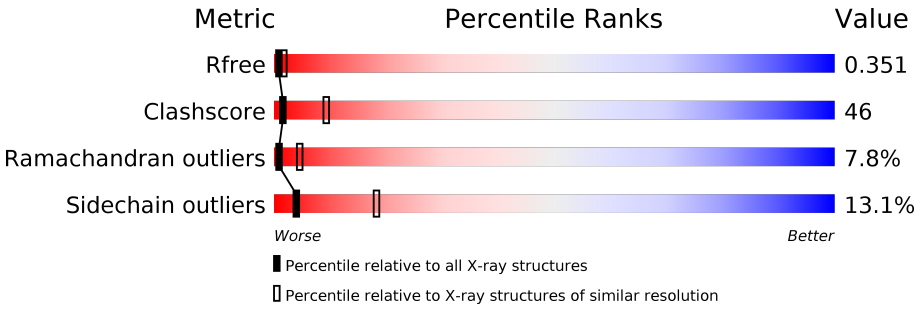
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.99 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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%

Mol	Chain	Length	Quality of chain
1	A	341	27% 41% 9% • 22%
1	B	341	28% 40% 9% • 22%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4097 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 Macrolide-specific efflux protein macA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	266	Total	C	N	O	S	0	0	0
			2043	1274	361	402	6			
1	B	267	Total	C	N	O	S	0	0	0
			2054	1280	365	403	6			

There are 6 discrepancies between the modelled and reference sequences:

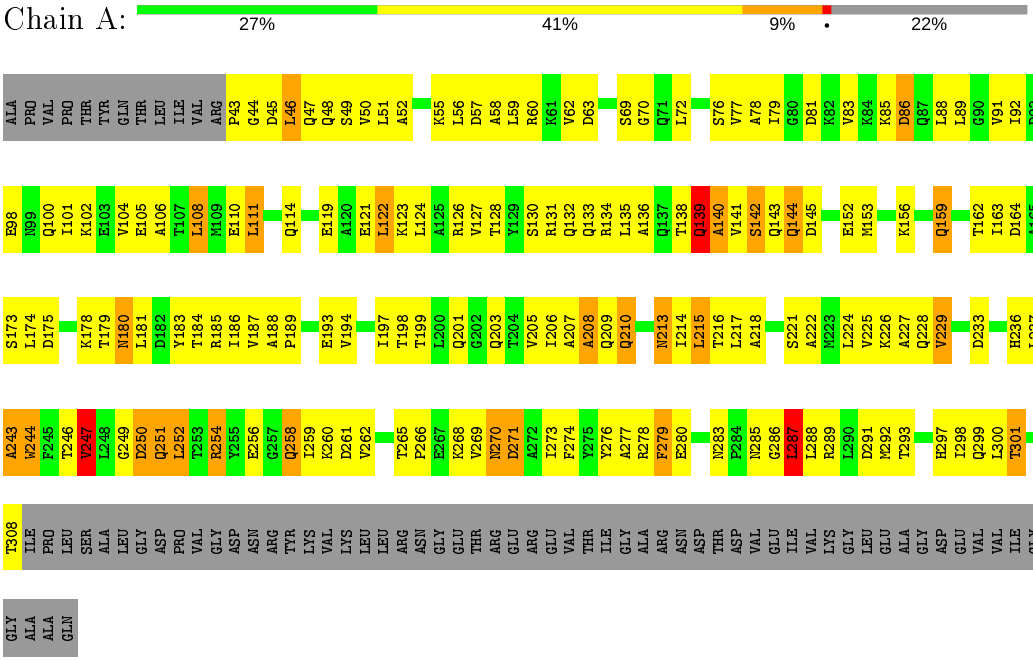
Chain	Residue	Modelled	Actual	Comment	Reference
A	139	GLN	LYS	SEE REMARK 999	UNP P75830
A	148	ASN	THR	SEE REMARK 999	UNP P75830
A	251	GLN	PRO	SEE REMARK 999	UNP P75830
B	139	GLN	LYS	SEE REMARK 999	UNP P75830
B	148	ASN	THR	SEE REMARK 999	UNP P75830
B	251	GLN	PRO	SEE REMARK 999	UNP P75830

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. 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. 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.

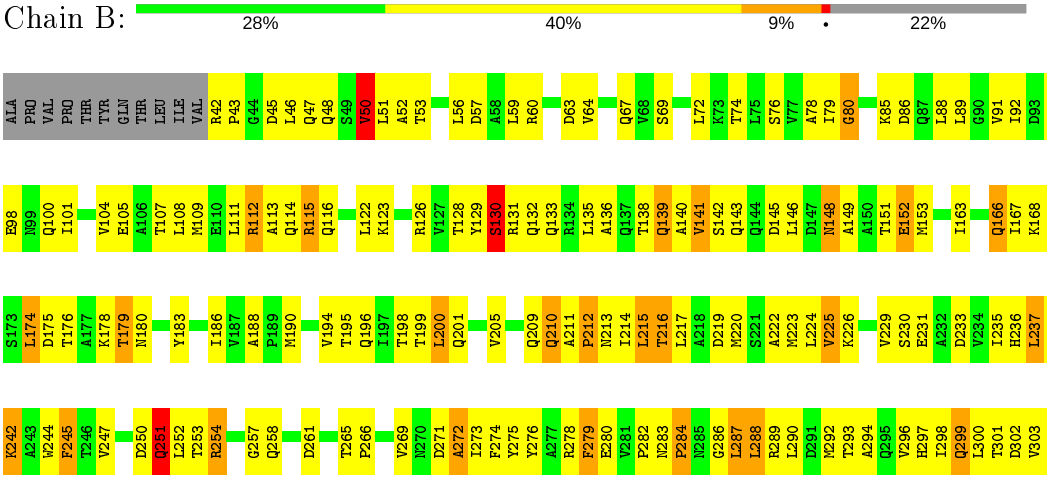
- Molecule 1: Macrolide-specific efflux protein macA

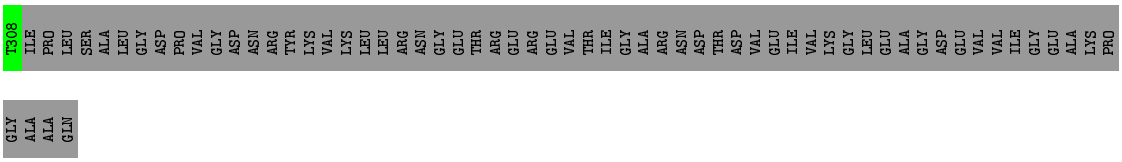
Chain A:



- Molecule 1: Macrolide-specific efflux protein macA

Chain B:




T308
ILE PRO LEU SER ALA LEU GLY ASP PRO VAL GLY ASP ASN ARG TYR LYS VAL LYS LEU LEU ARG ASN GLY GLU THR ARG GLU ARG GLU VAL THR ILE GLY ALA ARG ASN ASP THR ASP VAL GLU ILE VAL LYS GLY LEU GLU ALA GLY ASP GLU VAL VAL ILE GLY GLU ALA LYS PRO
GLY ALA ALA GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	128.52Å 128.52Å 110.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.69 – 2.99 49.69 – 3.00	Depositor EDS
% Data completeness (in resolution range)	92.0 (49.69-2.99) 92.1 (49.69-3.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 3.01Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.283 , 0.349 0.284 , 0.351	Depositor DCC
R_{free} test set	1950 reflections (9.81%)	wwPDB-VP
Wilson B-factor (Å^2)	37.7	Xtriage
Anisotropy	0.126	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , -6.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.447 for -h,-k,l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	4097	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.11% 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	0.45	0/2063	0.74	0/2797
1	B	0.44	0/2074	0.75	1/2812 (0.0%)
All	All	0.44	0/4137	0.74	1/5609 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	76	SER	N-CA-C	5.29	125.30	111.00

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	2043	0	2104	199	0
1	B	2054	0	2116	195	0
All	All	4097	0	4220	380	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 46.

The worst 5 of 380 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:159:GLN:HE21	1:A:159:GLN:HA	1.15	1.10
1:B:135:LEU:HD22	1:B:140:ALA:HB2	1.35	1.06
1:B:205:VAL:HG21	1:B:212:PRO:HG3	1.32	1.05
1:B:226:LYS:HZ2	1:B:278:ARG:HD3	1.23	1.04
1:A:285:ASN:ND2	1:A:286:GLY:H	1.55	1.03

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	264/341 (77%)	207 (78%)	39 (15%)	18 (7%)	1	6
1	B	265/341 (78%)	203 (77%)	39 (15%)	23 (9%)	1	3
All	All	529/682 (78%)	410 (78%)	78 (15%)	41 (8%)	1	4

5 of 41 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	140	ALA
1	A	221	SER
1	A	242	LYS
1	A	247	VAL
1	A	252	LEU

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	221/281 (79%)	189 (86%)	32 (14%)	3 15
1	B	222/281 (79%)	196 (88%)	26 (12%)	5 22
All	All	443/562 (79%)	385 (87%)	58 (13%)	4 18

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	283	ASN
1	B	50	VAL
1	B	279	PHE
1	A	287	LEU
1	A	301	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	299	GLN
1	B	100	GLN
1	B	251	GLN
1	B	96	GLN
1	B	114	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

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.