



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 05:57 pm BST

PDB ID : 2D45  
Title : Crystal structure of the MecI-mecA repressor-operator complex  
Authors : Safo, M.K.; Ko, T.-P.; Musayev, F.N.; Zhao, Q.; Wang, A.H.-J.; Archer, G.L.  
Deposited on : 2005-10-09  
Resolution : 3.80 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](mailto: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  
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

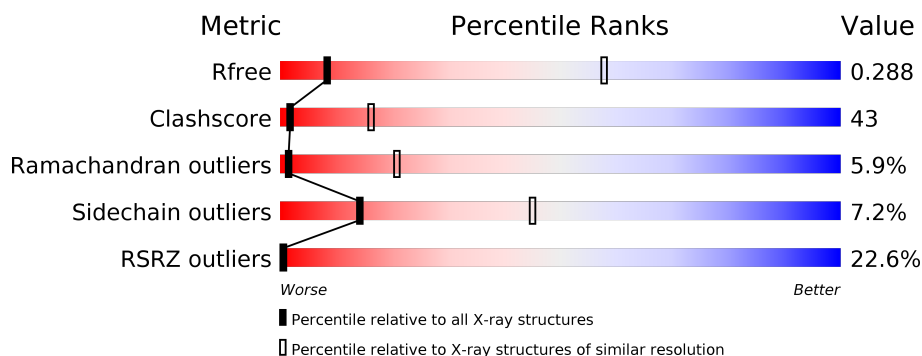
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.80 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	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

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  $\geq 3$ , 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	E	16	
1	F	16	
1	G	16	
1	H	16	
2	A	123	
2	B	123	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	123	<div><div></div><div>22%</div><div>39%</div><div>47%</div><div>9%</div><div>5%</div></div>
2	D	123	<div><div></div><div>21%</div><div>33%</div><div>53%</div><div>7%</div><div>7%</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5204 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 DNA chain called 5'-D(P\*TP\*AP\*CP\*TP\*AP\*CP\*AP\*TP\*AP\*TP\*GP\*TP\*AP\*GP\*TP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	16	Total	C	N	O	P	0	0	0
			328	158	58	96	16			
1	F	16	Total	C	N	O	P	0	0	0
			328	158	58	96	16			
1	G	13	Total	C	N	O	P	0	0	0
			265	128	46	78	13			
1	H	12	Total	C	N	O	P	0	0	0
			245	119	46	69	11			

- Molecule 2 is a protein called Methicillin resistance regulatory protein mecI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	117	Total	C	N	O	Se	0	0	0
			990	638	161	188	3			
2	B	119	Total	C	N	O	Se	0	0	0
			1007	648	165	191	3			
2	C	117	Total	C	N	O	Se	0	0	0
			990	638	161	188	3			
2	D	115	Total	C	N	O	Se	0	0	0
			971	625	159	184	3			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P68261
A	16	MSE	MET	MODIFIED RESIDUE	UNP P68261
A	21	MSE	MET	MODIFIED RESIDUE	UNP P68261
A	36	MSE	MET	MODIFIED RESIDUE	UNP P68261
B	1	MSE	MET	MODIFIED RESIDUE	UNP P68261
B	16	MSE	MET	MODIFIED RESIDUE	UNP P68261
B	21	MSE	MET	MODIFIED RESIDUE	UNP P68261

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	36	MSE	MET	MODIFIED RESIDUE	UNP P68261
C	1	MSE	MET	MODIFIED RESIDUE	UNP P68261
C	16	MSE	MET	MODIFIED RESIDUE	UNP P68261
C	21	MSE	MET	MODIFIED RESIDUE	UNP P68261
C	36	MSE	MET	MODIFIED RESIDUE	UNP P68261
D	1	MSE	MET	MODIFIED RESIDUE	UNP P68261
D	16	MSE	MET	MODIFIED RESIDUE	UNP P68261
D	21	MSE	MET	MODIFIED RESIDUE	UNP P68261
D	36	MSE	MET	MODIFIED RESIDUE	UNP P68261

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	1	Total O 1 1	0	0
3	F	1	Total O 1 1	0	0
3	G	4	Total O 4 4	0	0
3	H	5	Total O 5 5	0	0
3	A	16	Total O 16 16	0	0
3	B	14	Total O 14 14	0	0
3	C	11	Total O 11 11	0	0
3	D	28	Total O 28 28	0	0

### 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: 5'-D(P\*TP\*AP\*CP\*TP\*AP\*CP\*AP\*TP\*AP\*TP\*GP\*TP\*AP\*GP\*TP\*A)-3',

Chain E: 

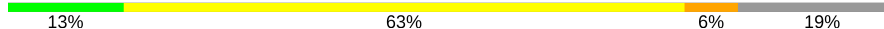


- Molecule 1: 5'-D(P\*TP\*AP\*CP\*TP\*AP\*CP\*AP\*TP\*AP\*TP\*GP\*TP\*AP\*GP\*TP\*A)-3',

Chain F: 

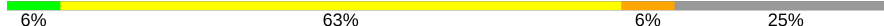


- Molecule 1: 5'-D(P\*TP\*AP\*CP\*TP\*AP\*CP\*AP\*TP\*AP\*TP\*GP\*TP\*AP\*GP\*TP\*A)-3',

Chain G: 



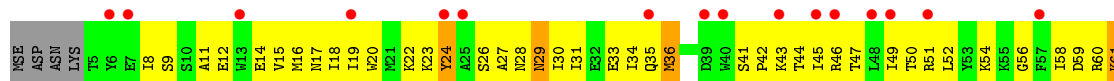
- Molecule 1: 5'-D(P\*TP\*AP\*CP\*TP\*AP\*CP\*AP\*TP\*AP\*TP\*GP\*TP\*AP\*GP\*TP\*A)-3',

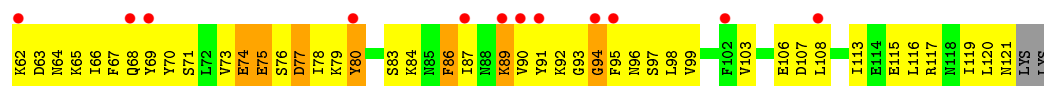
Chain H: 



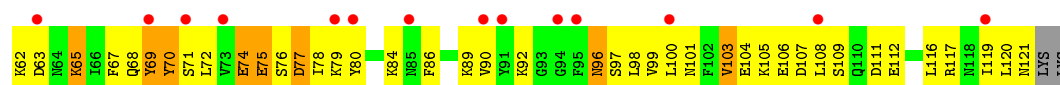
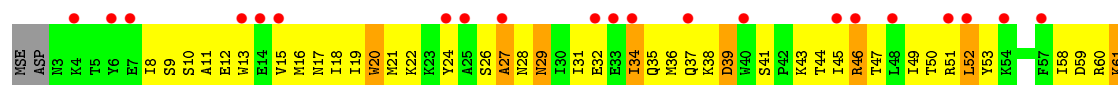
- Molecule 2: Methicillin resistance regulatory protein mecI

Chain A: 

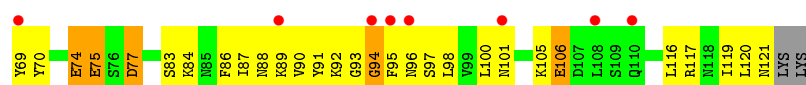
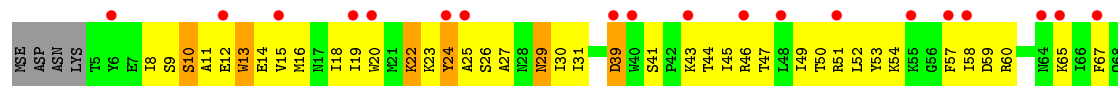




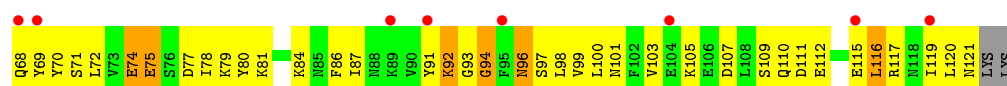
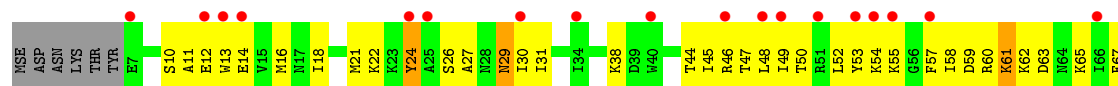
- Molecule 2: Methicillin resistance regulatory protein mecI



- Molecule 2: Methicillin resistance regulatory protein mecI



- Molecule 2: Methicillin resistance regulatory protein mecI



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.46Å 61.46Å 419.78Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 3.80 39.81 – 0.58	Depositor EDS
% Data completeness (in resolution range)	97.6 (40.00-3.80) 0.4 (39.81-0.58)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.264 , 0.297 0.253 , 0.288	Depositor DCC
$R_{free}$ test set	493 reflections (5.37%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	(Not available)	Xtriage
Anisotropy	(Not available)	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 88.0	EDS
L-test for twinning <sup>1</sup>	$\langle  L  \rangle =$ (Not available), $\langle L^2 \rangle =$ (Not available)	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	5204	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	118.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: (*Not available*)

<sup>1</sup>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	E	1.15	2/367 (0.5%)	0.88	0/564
1	F	0.93	0/367	0.82	0/564
1	G	0.72	0/296	0.82	0/454
1	H	0.88	0/275	0.71	0/423
2	A	0.94	0/1005	0.87	0/1344
2	B	1.05	0/1022	0.91	0/1366
2	C	0.91	1/1005 (0.1%)	0.82	0/1344
2	D	0.81	0/985	0.82	0/1316
All	All	0.94	3/5322 (0.1%)	0.84	0/7375

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	E	0	1
1	F	0	2
1	G	0	1
1	H	0	1
2	A	0	1
All	All	0	6

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	212	DT	N1-C2	-7.20	1.32	1.38
2	C	13	TRP	CB-CG	-7.10	1.37	1.50
1	E	213	DA	C5-C6	-6.05	1.35	1.41

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	212	DT	Sidechain
1	F	211	DG	Sidechain
1	F	212	DT	Sidechain
1	G	211	DG	Sidechain
1	H	212	DT	Sidechain

## 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	E	328	0	183	25	0
1	F	328	0	183	18	1
1	G	265	0	149	26	0
1	H	245	0	138	21	0
2	A	990	0	997	101	1
2	B	1007	0	1016	115	0
2	C	990	0	997	95	1
2	D	971	0	981	90	0
3	A	16	0	0	0	0
3	B	14	0	0	0	0
3	C	11	0	0	0	0
3	D	28	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	4	0	0	0	0
3	H	5	0	0	0	0
All	All	5204	0	4644	421	2

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 43.

The worst 5 of 421 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:C:16:MSE:HE1	2:C:57:PHE:HB3	1.22	1.13

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:92:LYS:HE2	2:B:92:LYS:HA	1.31	1.09
1:H:211:DG:H2''	1:H:212:DT:H5''	1.33	1.07
2:C:16:MSE:HE1	2:C:57:PHE:CB	1.85	1.07
2:A:74:GLU:HB3	2:A:77:ASP:HB2	1.44	0.99

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:36:MSE:O	2:C:46:ARG:NH2[1_665]	2.16	0.04
1:F:201:DT:P	1:F:216:DA:O3'[4_765]	2.17	0.03

## 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	
2	A	115/123 (94%)	95 (83%)	11 (10%)	9 (8%)	1	15
2	B	117/123 (95%)	91 (78%)	20 (17%)	6 (5%)	2	23
2	C	115/123 (94%)	87 (76%)	22 (19%)	6 (5%)	2	23
2	D	113/123 (92%)	85 (75%)	22 (20%)	6 (5%)	2	23
All	All	460/492 (94%)	358 (78%)	75 (16%)	27 (6%)	1	21

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	75	GLU
2	C	75	GLU
2	C	94	GLY
2	D	75	GLU
2	A	35	GLN

### 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	
2	A	111/113 (98%)	105 (95%)	6 (5%)	22	53
2	B	113/113 (100%)	98 (87%)	15 (13%)	4	22
2	C	111/113 (98%)	105 (95%)	6 (5%)	22	53
2	D	109/113 (96%)	104 (95%)	5 (5%)	27	56
All	All	444/452 (98%)	412 (93%)	32 (7%)	14	45

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	65	LYS
2	B	74	GLU
2	D	74	GLU
2	B	70	TYR
2	B	77	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	29	ASN
2	B	37	GLN
2	C	85	ASN
2	A	68	GLN
2	C	29	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, 95<sup>th</sup> 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(Å²)	Q<0.9	
1	E	16/16 (100%)	-0.54	0	100	100	97, 117, 141, 144	0
1	F	16/16 (100%)	-0.49	0	100	100	109, 124, 135, 141	0
1	G	13/16 (81%)	-0.29	0	100	100	127, 141, 162, 163	0
1	H	12/16 (75%)	-0.37	0	100	100	122, 139, 182, 191	0
2	A	114/123 (92%)	1.28	28 (24%)	0	0	71, 97, 173, 175	0
2	B	116/123 (94%)	1.54	35 (30%)	0	0	74, 102, 141, 146	0
2	C	114/123 (92%)	1.18	27 (23%)	0	0	76, 124, 183, 186	0
2	D	112/123 (91%)	0.97	26 (23%)	0	0	87, 123, 152, 154	0
All	All	513/556 (92%)	1.06	116 (22%)	0	0	71, 114, 171, 191	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	40	TRP	13.5
2	A	69	TYR	7.7
2	B	37	GLN	7.5
2	D	12	GLU	6.8
2	A	24	TYR	6.6

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.