



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 01:32 pm GMT

PDB ID : 4R1I  
Title : Structure and Function of Neisseria gonorrhoeae MtrF Illuminates a Class of Antimetabolite Efflux Pumps  
Authors : Su, C.-C.; Bolla, J.R.; Yu, E.W.  
Deposited on : 2014-08-06  
Resolution : 3.96 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

<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

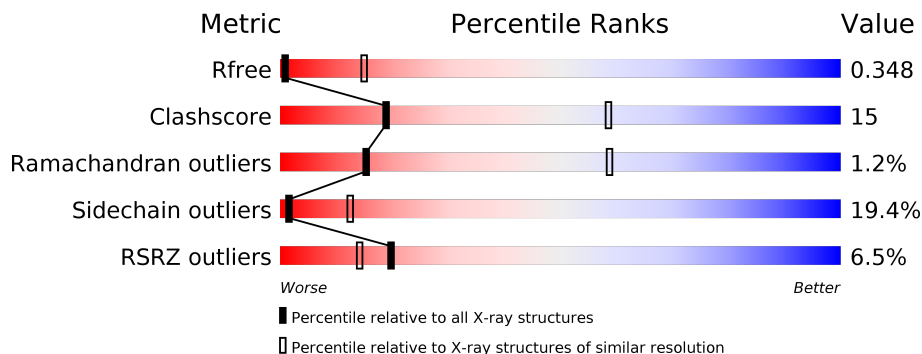
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.96 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	1057 (4.32-3.60)
Clashscore	112137	1154 (4.32-3.60)
Ramachandran outliers	110173	1111 (4.32-3.60)
Sidechain outliers	110143	1101 (4.32-3.60)
RSRZ outliers	101464	1069 (4.32-3.60)

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. 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	522	<div> <div>7%</div> <div> <div></div> <div>55%</div> <div>36%</div> <div>7%</div> <div>.</div> </div> </div>
1	B	522	<div> <div>5%</div> <div> <div></div> <div>56%</div> <div>36%</div> <div>5%</div> <div>.</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7684 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 Aminobenzoyl-glutamate transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	507	Total	C	N	O	S	0	0	0
			3846	2566	609	654	17			
1	B	506	Total	C	N	O	S	0	0	0
			3838	2563	605	653	17			

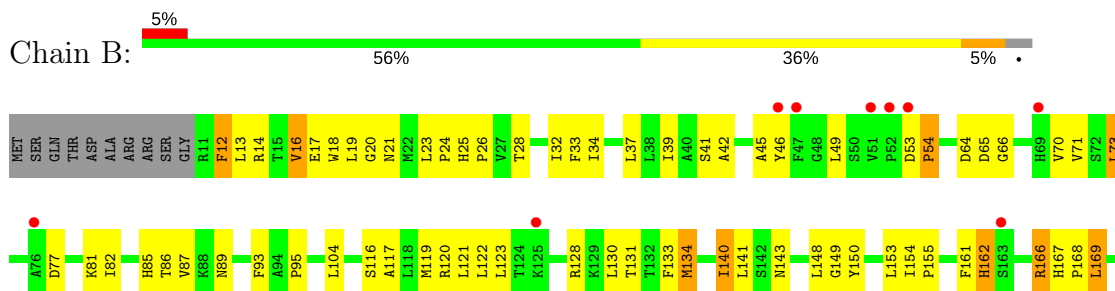
### 3 Residue-property plots [i](#)

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 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: Aminobenzoyl-glutamate transporter



#### • Molecule 1: Aminobenzoyl-glutamate transporter



Y511	G512	P516	THR	LEU	TYR	PRO	ALA	PRO	E439	V440	I441	Q442	A443	A444	Y445	R446	I447	V451	T452	I455	T456	P457	M458	M459	S460	Y461	F462	G463	L464	I465	M466	V469	M470	K471	Y472	K473	A476	G477	V478	L481	I482	S483	L486	P487	Y488	F492	W496	L499	F500	C501	I502	W503	V504	F505	V506	L507	G508	L509	P510	G339	E340	R341	E342	V343	V344	N345	A346	M347	A348	S352	L358	I361	F362	Q366	F367	F370	F371	N372	W373	T374	N375	Q378	Y379	L388	L398	F399	A407	F408	I409	N410	L411	M412	I413	G414	S415	A416	Q419	W420	A421	V422	V428	P429	G331	M430	L431	M432	A437	P438	E256	K257	D258	H261	P267	L268	E269	Y270	K271	Q272	L273	I274	W275	A276	V279	A282	L283	S284	A285	L286	L287	A288	W289	S290	I291	A294	L298	R299	L305	V306	F310	F311	L312	K313	S314	I315	V316	L321	L325	I328	V329	Y330	Q349	R332	L333	S336	I337	R338	L172	G182	Y183	S184	A185	N186	L187	F188	L189	I192	D193	L196	I199	T200	Q201	Q202	I206	L207	D210	Y211	V212	V213	A217	N218	F221	S225	T226	I229	I232	G233	Y234	F235	T240	V241	E242	P243	Q244	L245	G246	P247	Y248	Q249	S250	D251	L252	S253	Q254	E255
------	------	------	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	120.77Å 120.77Å 233.90Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.74 – 3.96 47.74 – 3.96	Depositor EDS
% Data completeness (in resolution range)	97.4 (47.74-3.96) 97.8 (47.74-3.96)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 4.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.311 , 0.338 0.298 , 0.348	Depositor DCC
$R_{free}$ test set	868 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	202.7	Xtriage
Anisotropy	0.407	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 50.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.368 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	7684	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	163.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.24% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</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	A	0.27	0/3945	0.49	0/5381
1	B	0.25	0/3938	0.48	0/5374
All	All	0.26	0/7883	0.48	0/10755

There are no bond length outliers.

There are no bond angle outliers.

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	3846	0	3998	129	0
1	B	3838	0	3990	106	0
All	All	7684	0	7988	231	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 15.

The worst 5 of 231 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:437:ALA:H	1:A:512:GLY:H	1.22	0.88
1:A:456:THR:HG23	1:A:482:ILE:HD12	1.60	0.82

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:393:LEU:HB3	1:A:398:LEU:HB2	1.65	0.79
1:B:131:THR:HA	1:B:134:MET:HB2	1.67	0.76
1:A:184:SER:HB2	1:A:446:ARG:HG2	1.69	0.75

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	505/522 (97%)	453 (90%)	46 (9%)	6 (1%)	15	59
1	B	504/522 (97%)	447 (89%)	51 (10%)	6 (1%)	15	59
All	All	1009/1044 (97%)	900 (89%)	97 (10%)	12 (1%)	15	59

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	209	PRO
1	A	243	PRO
1	B	217	ALA
1	A	250	SER
1	A	510	PRO

### 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	403/418 (96%)	326 (81%)	77 (19%)	2	13
1	B	403/418 (96%)	324 (80%)	79 (20%)	1	12
All	All	806/836 (96%)	650 (81%)	156 (19%)	1	12

5 of 156 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	496	TRP
1	B	65	ASP
1	B	442	GLN
1	A	501	CYS
1	B	14	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	25	HIS
1	A	167	HIS
1	A	410	ASN
1	A	419	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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 [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(Å <sup>2</sup> )	Q < 0.9
1	A	507/522 (97%)	-0.20	38 (7%)	15 12	54, 157, 245, 345	0
1	B	506/522 (96%)	-0.25	28 (5%)	26 20	44, 153, 257, 321	0
All	All	1013/1044 (97%)	-0.22	66 (6%)	20 14	44, 155, 254, 345	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	478	VAL	8.0
1	A	297	ILE	7.3
1	B	399	PHE	7.3
1	B	477	GLY	7.2
1	B	478	VAL	6.3

### 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.