



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 05:31 pm BST

PDB ID : 5WIC  
Title : TraE protein in complex with 2-Furoic Acid (FOA)  
Authors : Casu, B.; Arya, T.; Bessette, B.; Baron, C.  
Deposited on : 2017-07-19  
Resolution : 2.55 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

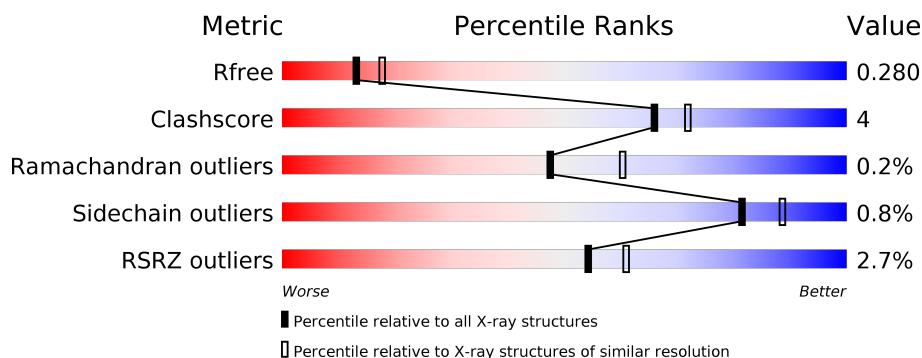
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	A	163	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 82%, yellow 82%, yellow 88%, grey 88%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>82%</span> <span>6%</span> <span>12%</span> </div> </div>
1	B	163	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 80%, yellow 80%, yellow 87%, grey 87%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>80%</span> <span>7%</span> <span>13%</span> </div> </div>
1	C	163	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 3%, green 3%, green 77%, yellow 77%, yellow 84%, grey 84%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>77%</span> <span>7%</span> <span>16%</span> </div> </div>
1	D	163	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 6%, green 6%, green 69%, yellow 69%, yellow 84%, grey 84%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>69%</span> <span>15%</span> <span>16%</span> </div> </div>

## 2 Entry composition [i](#)

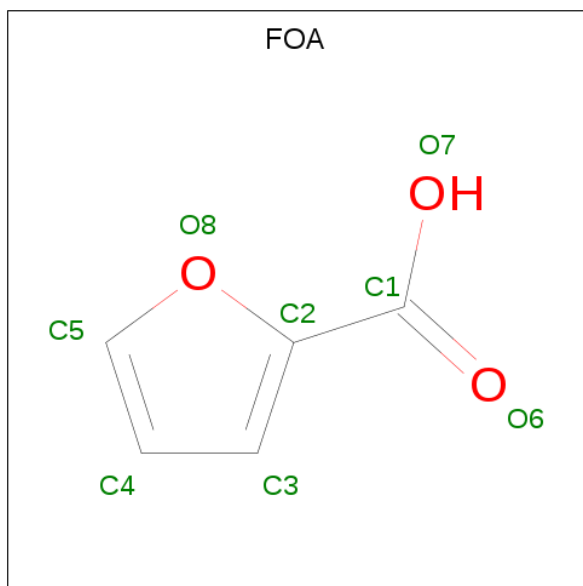
There are 3 unique types of molecules in this entry. The entry contains 9146 atoms, of which 4463 are hydrogens and 0 are deuteriums.

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 Conjugal transfer protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	143	Total	C	H	N	O	S	0	0	0
			2307	736	1139	206	223	3			
1	B	142	Total	C	H	N	O	S	0	0	0
			2294	732	1133	205	221	3			
1	C	137	Total	C	H	N	O	S	0	0	0
			2215	710	1095	196	211	3			
1	D	137	Total	C	H	N	O	S	0	0	0
			2205	707	1090	194	211	3			

- Molecule 2 is 2-FUROIC ACID (three-letter code: FOA) (formula: C<sub>5</sub>H<sub>4</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	H	O	0	0
			11	5	3	3		
2	B	1	Total	C	H	O	0	0
			11	5	3	3		


- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	46	Total 46	O 46	0	0
3	B	43	Total 43	O 43	0	0
3	C	7	Total 7	O 7	0	0
3	D	7	Total 7	O 7	0	0

### 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: Conjugal transfer protein

Chain A: 




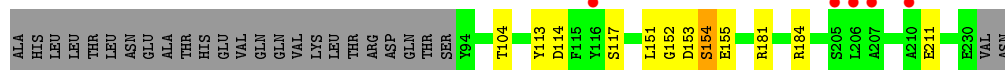
- Molecule 1: Conjugal transfer protein

Chain B: 



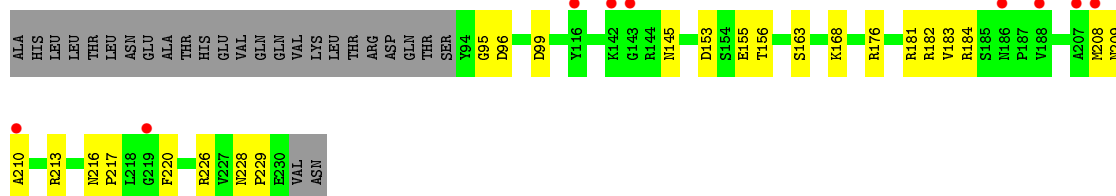
- Molecule 1: Conjugal transfer protein

Chain C: 



- Molecule 1: Conjugal transfer protein

Chain D: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.32Å 123.77Å 108.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.77 – 2.55 41.38 – 2.17	Depositor EDS
% Data completeness (in resolution range)	97.8 (38.77-2.55) 88.7 (41.38-2.17)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.00 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.225 , 0.281 0.231 , 0.280	Depositor DCC
$R_{free}$ test set	2000 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.9	Xtriage
Anisotropy	0.482	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 35.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	9146	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.25% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: FOA

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	0.48	0/1194	0.46	0/1619
1	B	0.49	0/1187	0.48	0/1609
1	C	0.44	0/1146	0.47	0/1554
1	D	0.42	0/1140	0.45	0/1546
All	All	0.46	0/4667	0.47	0/6328

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 [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	1168	1139	1139	8	0
1	B	1161	1133	1132	8	0
1	C	1120	1095	1095	6	0
1	D	1115	1090	1090	15	0
2	B	16	6	6	1	0
3	A	46	0	0	6	0
3	B	43	0	0	3	0
3	C	7	0	0	1	0
3	D	7	0	0	1	0
All	All	4683	4463	4462	36	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 4.

All (36) 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:C:151:LEU:O	1:C:154:SER:OG	1.76	1.02
1:D:95:GLY:H	1:D:168:LYS:HZ1	1.19	0.85
1:C:153:ASP:OD1	1:C:184:ARG:NH2	2.14	0.79
1:C:104:THR:OG1	3:C:301:HOH:O	1.99	0.78
1:D:213:ARG:NH2	1:D:217:PRO:O	2.18	0.76
1:D:95:GLY:H	1:D:168:LYS:NZ	1.83	0.76
1:B:150:VAL:O	3:B:401:HOH:O	2.06	0.72
1:D:216:ASN:ND2	1:D:220:PHE:O	2.22	0.70
1:D:182:ARG:NH1	1:D:183:VAL:O	2.25	0.69
1:A:190:ASP:OD1	3:A:301:HOH:O	2.11	0.68
1:B:202:GLU:OE2	1:B:221:ARG:NE	2.26	0.67
1:A:202:GLU:OE2	1:A:221:ARG:NE	2.29	0.62
1:B:230:GLU:OE1	3:B:402:HOH:O	2.16	0.59
1:D:163:SER:OG	1:D:176:ARG:NH1	2.37	0.58
1:D:96:ASP:OD2	3:D:302:HOH:O	2.17	0.57
1:A:194:ARG:NH1	3:A:309:HOH:O	2.39	0.56
1:A:211:GLU:HG2	3:A:321:HOH:O	2.05	0.55
1:A:137:TYR:OH	3:A:302:HOH:O	2.19	0.51
1:D:216:ASN:HD21	1:D:220:PHE:N	2.08	0.51
1:D:155:GLU:OE1	1:D:181:ARG:NH2	2.41	0.51
1:B:203:TYR:O	3:B:403:HOH:O	2.18	0.50
1:A:184:ARG:NH1	1:B:170:HIS:NE2	2.62	0.48
1:D:95:GLY:N	1:D:168:LYS:HZ1	1.99	0.48
1:B:113:TYR:HD1	2:B:302:FOA:H3	1.81	0.46
1:D:209:ASN:OD1	1:D:210:ALA:N	2.48	0.46
1:B:209:ASN:OD1	1:B:212:GLN:HG3	2.16	0.45
1:B:102:TRP:CH2	1:B:217:PRO:HD2	2.53	0.44
1:D:228:ASN:HA	1:D:229:PRO:HD3	1.84	0.43
1:C:114:ASP:HB3	1:C:117:SER:OG	2.19	0.43
1:D:153:ASP:OD2	1:D:184:ARG:NE	2.50	0.42
1:C:155:GLU:OE1	1:C:181:ARG:NH2	2.51	0.42
1:A:105:GLN:NE2	3:A:315:HOH:O	2.52	0.42
1:A:176:ARG:NH1	3:A:308:HOH:O	2.37	0.41
1:D:99:ASP:OD2	1:D:168:LYS:HD3	2.20	0.41
1:D:156:THR:O	1:D:182:ARG:N	2.51	0.41
1:C:113:TYR:OH	1:C:152:GLY:O	2.29	0.40

There are no symmetry-related clashes.



## 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	
1	A	141/163 (86%)	137 (97%)	4 (3%)	0	100	100
1	B	140/163 (86%)	139 (99%)	1 (1%)	0	100	100
1	C	135/163 (83%)	128 (95%)	6 (4%)	1 (1%)	22	30
1	D	135/163 (83%)	127 (94%)	8 (6%)	0	100	100
All	All	551/652 (84%)	531 (96%)	19 (3%)	1 (0%)	47	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	211	GLU

### 5.3.2 Protein sidechains

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	129/147 (88%)	129 (100%)	0	100	100
1	B	128/147 (87%)	128 (100%)	0	100	100
1	C	123/147 (84%)	122 (99%)	1 (1%)	81	88
1	D	122/147 (83%)	119 (98%)	3 (2%)	47	62
All	All	502/588 (85%)	498 (99%)	4 (1%)	81	88

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	154	SER
1	D	145	ASN
1	D	208	MET
1	D	226	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	FOA	B	301	-	1,8,8	0.11	0	1,10,10	1.42	0
2	FOA	B	302	-	1,8,8	0.13	0	1,10,10	1.34	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	FOA	B	301	-	-	0/0/4/4	0/1/1/1
2	FOA	B	302	-	-	0/0/4/4	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	302	FOA	1	0

## 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 ⓘ

### 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<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	143/163 (87%)	-0.24	0 100 100	32, 43, 69, 83	0
1	B	142/163 (87%)	-0.11	1 (0%) 87 90	32, 45, 69, 81	0
1	C	137/163 (84%)	0.29	5 (3%) 42 49	48, 74, 95, 115	0
1	D	137/163 (84%)	0.43	9 (6%) 18 21	49, 78, 106, 132	0
All	All	559/652 (85%)	0.09	15 (2%) 54 61	32, 62, 96, 132	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	210	ALA	7.9
1	C	206	LEU	6.3
1	C	207	ALA	5.9
1	D	143	GLY	5.7
1	D	210	ALA	3.2
1	D	207	ALA	3.0
1	D	208	MET	2.8
1	D	142	LYS	2.6
1	D	116	TYR	2.6
1	C	116	TYR	2.6
1	C	205	SER	2.5
1	B	116	TYR	2.4
1	D	186	ASN	2.3
1	D	188	VAL	2.3
1	D	219	GLY	2.1

### 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 [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	FOA	B	302	8/8	0.75	0.22	38,44,52,57	10
2	FOA	B	301	8/8	0.93	0.14	56,59,69,70	3

### 6.5 Other polymers [i](#)

There are no such residues in this entry.