



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2016 – 08:32 PM EDT

PDB ID : 5LJI  
Title : Streptococcus pneumonia TIGR4 flavodoxin: structural and biophysical characterization of a novel drug target  
Authors : Rodriguez-Cardenas, A.; Rojas, A.L.; Velazquez-Campoy, A.; Hurtado-Guerrero, R.; Sancho, J.  
Deposited on : 2016-07-18  
Resolution : 2.07 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

<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:

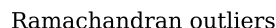
MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027939



**i**

## X-RAY DIFFRACTION

A.

Ramachandran outliers

electron density. The numeric value is given above the bar.

1

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:



Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CA	A	203	-	-	-	X
3	CA	C	203	-	-	-	X
3	CA	C	206	-	-	-	X



## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2359 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.

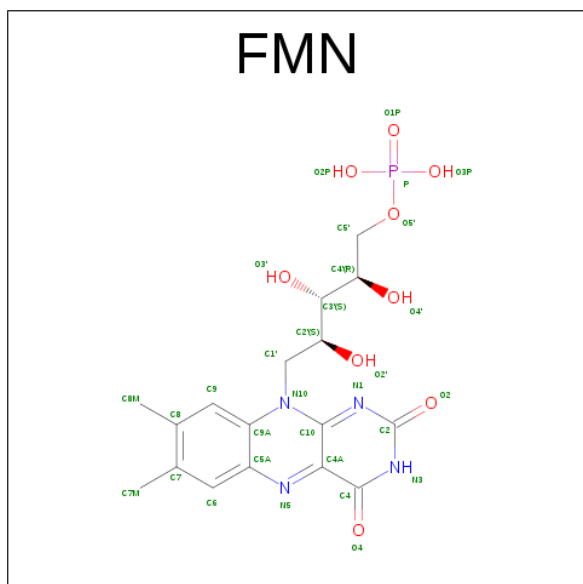
- Molecule 1 is a protein called Flavodoxin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	145	Total	C	N	O	S	Se	0	1	0
			1112	696	165	244	4	3			
1	C	145	Total	C	N	O	S	Se	0	0	0
			1109	694	165	244	3	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	148	ALA	-	expression tag	UNP A0A0B7M5W3
C	148	ALA	-	expression tag	UNP A0A0B7M5W3

- Molecule 2 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula: C<sub>17</sub>H<sub>21</sub>N<sub>4</sub>O<sub>9</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Ca	0	0
			2	2		
3	C	5	Total	Ca	0	0
			5	5		

- Molecule 4 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	35	Total	O	0	0
			35	35		
4	C	34	Total	O	0	0
			34	34		



### 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 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.

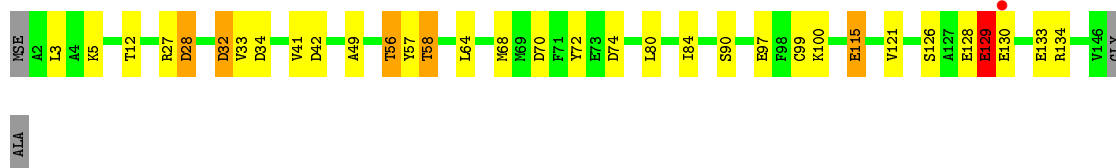
#### • Molecule 1: Flavodoxin

Chain A: 



#### • Molecule 1: Flavodoxin

Chain C: 



ALA



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.68 Å   101.68 Å   68.28 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	37.00 – 2.07 37.00 – 2.07	Depositor EDS
% Data completeness (in resolution range)	99.2 (37.00-2.07) 99.2 (37.00-2.07)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.31 (at 2.06 Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.187   ,   0.229 0.200   ,   0.234	Depositor DCC
$R_{free}$ test set	799 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	28.1	Xtriage
Anisotropy	0.431	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36   ,   29.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.478 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2359	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.14% 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

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

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	0.89	1/1125 (0.1%)	1.04	6/1518 (0.4%)
1	C	0.90	1/1119 (0.1%)	1.03	4/1510 (0.3%)
All	All	0.90	2/2244 (0.1%)	1.03	10/3028 (0.3%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	34	ASP	CB-CG	5.42	1.63	1.51
1	A	34	ASP	CB-CG	5.23	1.62	1.51

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	28	ASP	CB-CG-OD1	7.24	124.82	118.30
1	A	28	ASP	CB-CG-OD1	7.03	124.63	118.30
1	A	56	THR	CB-CA-C	-6.93	92.90	111.60
1	C	56	THR	CB-CA-C	-6.78	93.29	111.60
1	A	28	ASP	CB-CG-OD2	-6.03	112.87	118.30
1	C	129	GLU	N-CA-C	-5.87	95.16	111.00
1	A	74	ASP	CB-CG-OD1	5.43	123.19	118.30
1	A	32	ASP	CB-CG-OD2	-5.16	113.65	118.30
1	C	74	ASP	CB-CG-OD1	5.08	122.87	118.30
1	A	103	ASP	CB-CG-OD1	5.01	122.81	118.30

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	1112	0	1043	12	0
1	C	1109	0	1038	15	0
2	A	31	0	19	0	0
2	C	31	0	19	1	0
3	A	2	0	0	0	1
3	C	5	0	0	0	1
4	A	35	0	0	4	0
4	C	34	0	0	4	0
All	All	2359	0	2119	27	1

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 6.

All (27) 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:5:LYS:HD2	1:A:49:ALA:HB2	1.67	0.76
1:C:97:GLU:HG2	4:C:322:HOH:O	1.85	0.76
1:C:5:LYS:HD2	1:C:49:ALA:HB2	1.69	0.74
1:C:58:THR:HG22	4:C:316:HOH:O	1.99	0.62
1:A:97:GLU:HG2	4:A:326:HOH:O	2.00	0.60
1:A:97:GLU:HB3	1:A:100:LYS:HG3	1.83	0.58
1:C:97:GLU:HB3	1:C:100:LYS:HG3	1.83	0.58
1:A:58:THR:HG22	4:A:314:HOH:O	2.07	0.55
1:C:58:THR:CG2	1:C:99:CYS:HA	2.37	0.55
1:A:58:THR:CG2	1:A:99:CYS:HA	2.37	0.54
1:C:70:ASP:HB2	4:C:321:HOH:O	2.07	0.53
1:A:64:LEU:HD21	1:A:72:TYR:CG	2.46	0.51
1:C:64:LEU:HD21	1:C:72:TYR:CG	2.46	0.51
1:C:3:LEU:HD13	1:C:32:ASP:HB3	1.93	0.50
1:A:134:ARG:NH1	4:A:302:HOH:O	2.34	0.50
1:A:3:LEU:HD13	1:A:32:ASP:HB3	1.95	0.48
1:C:134:ARG:HD2	4:C:304:HOH:O	2.15	0.46
1:A:57:TYR:HA	1:A:90:SER:O	2.16	0.46
1:C:12:THR:OG1	2:C:201:FMN:O1P	2.28	0.46

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:57:TYR:HA	1:C:90:SER:O	2.17	0.45
1:C:128:GLU:CG	1:C:129:GLU:N	2.79	0.45
1:A:84:ILE:HD13	1:A:115:GLU:HG2	2.00	0.44
1:C:84:ILE:HD13	1:C:115:GLU:HG2	2.00	0.43
1:A:134:ARG:HD2	4:A:302:HOH:O	2.19	0.41
1:C:3:LEU:CD1	1:C:32:ASP:HB3	2.50	0.41
1:C:128:GLU:CG	1:C:129:GLU:H	2.34	0.41
1:A:58:THR:HG21	1:A:101:ALA:H	1.86	0.40

All (1) 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)
3:A:203:CA:CA	3:C:206:CA:CA[9_545]	0.83	1.37

## 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	144/148 (97%)	142 (99%)	2 (1%)	0	100	100
1	C	143/148 (97%)	142 (99%)	1 (1%)	0	100	100
All	All	287/296 (97%)	284 (99%)	3 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	120/116 (103%)	106 (88%)	14 (12%)	7	2
1	C	119/116 (103%)	103 (87%)	16 (13%)	5	1
All	All	239/232 (103%)	209 (87%)	30 (13%)	5	1

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	ARG
1	A	28	ASP
1	A	32	ASP
1	A	33	VAL
1	A	41	VAL
1	A	42	ASP
1	A	56	THR
1	A	58	THR
1	A	68	MSE
1	A	70	ASP
1	A	80	LEU
1	A	115	GLU
1	A	121	VAL
1	A	128	GLU
1	C	27	ARG
1	C	28	ASP
1	C	32	ASP
1	C	33	VAL
1	C	41	VAL
1	C	42	ASP
1	C	56	THR
1	C	58	THR
1	C	68	MSE
1	C	80	LEU
1	C	115	GLU
1	C	121	VAL
1	C	126	SER
1	C	129	GLU
1	C	130	GLU
1	C	133	GLU

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 ⓘ

Of 9 ligands modelled in this entry, 7 are monoatomic - leaving 2 for Mogul analysis.

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	FMN	A	201	-	32,33,33	1.47	6 (18%)	34,50,50	3.20	9 (26%)
2	FMN	C	201	-	32,33,33	1.48	5 (15%)	34,50,50	3.03	10 (29%)

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	FMN	A	201	-	-	0/18/18/18	0/3/3/3
2	FMN	C	201	-	-	0/18/18/18	0/3/3/3

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	201	FMN	C6-C5A	-3.15	1.36	1.41

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	201	FMN	C2-N1	-2.78	1.32	1.38
2	A	201	FMN	C2-N1	-2.36	1.33	1.38
2	C	201	FMN	C6-C5A	-2.11	1.38	1.41
2	A	201	FMN	C4-C4A	2.03	1.45	1.41
2	C	201	FMN	C8-C7	2.88	1.48	1.41
2	A	201	FMN	C8-C7	2.91	1.48	1.41
2	A	201	FMN	C4A-C10	3.50	1.47	1.40
2	C	201	FMN	C4A-C10	3.52	1.47	1.40
2	A	201	FMN	C9A-C5A	3.64	1.50	1.42
2	C	201	FMN	C9A-C5A	4.09	1.51	1.42

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	201	FMN	C4-C4A-C10	-9.51	113.86	119.94
2	C	201	FMN	C4-C4A-C10	-9.31	113.98	119.94
2	C	201	FMN	N3-C2-N1	-5.27	118.81	127.69
2	A	201	FMN	N3-C2-N1	-5.17	118.98	127.69
2	A	201	FMN	C4A-C4-N3	-3.67	118.73	123.52
2	A	201	FMN	C4A-C10-N10	-3.40	118.05	120.52
2	C	201	FMN	C4A-C10-N10	-3.11	118.26	120.52
2	C	201	FMN	C4A-C4-N3	-3.10	119.47	123.52
2	C	201	FMN	O3'-C3'-C2'	-3.02	100.89	108.73
2	A	201	FMN	O3'-C3'-C2'	-2.77	101.54	108.73
2	A	201	FMN	C5A-C9A-N10	2.41	119.38	117.58
2	C	201	FMN	O3P-P-O2P	2.53	116.73	107.44
2	C	201	FMN	C4-C4A-N5	3.00	122.34	118.70
2	C	201	FMN	C1'-N10-C9A	3.08	122.40	118.83
2	C	201	FMN	C5A-C9A-N10	3.52	120.22	117.58
2	A	201	FMN	C4-C4A-N5	3.61	123.08	118.70
2	A	201	FMN	C1'-N10-C9A	4.14	123.63	118.83
2	C	201	FMN	C4-N3-C2	10.71	124.09	115.16
2	A	201	FMN	C4-N3-C2	11.83	125.03	115.16

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	C	201	FMN	1	0



## 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 [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	142/148 (95%)	-0.44	0 100 100	24, 37, 56, 70	0
1	C	142/148 (95%)	-0.44	1 (0%) 89 90	23, 37, 54, 76	0
All	All	284/296 (95%)	-0.44	1 (0%) 93 94	23, 37, 56, 76	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	130	GLU	2.2

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	CA	C	203	1/1	0.99	0.27	16.39	40,40,40,40	0
3	CA	C	206	1/1	1.00	0.18	4.12	52,52,52,52	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CA	A	203	1/1	1.00	0.18	2.47	48,48,48,48	0
2	FMN	C	201	31/31	0.97	0.09	-0.27	25,30,47,49	0
2	FMN	A	201	31/31	0.97	0.08	-0.32	23,30,43,47	0
3	CA	C	202	1/1	0.99	0.04	-1.56	34,34,34,34	0
3	CA	A	202	1/1	0.99	0.04	-2.37	34,34,34,34	0
3	CA	C	204	1/1	0.99	0.08	-	31,31,31,31	0
3	CA	C	205	1/1	0.98	0.06	-	35,35,35,35	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.